

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 9/48, 15/57, C07K 16/40, C12N 15/10, C12Q 1/68, G01N 33/53, A61K 38/48		A1	(11) International Publication Number: WO 00/70032 (43) International Publication Date: 23 November 2000 (23.11.00)
(21) International Application Number: PCT/US00/11932 (22) International Filing Date: 3 May 2000 (03.05.00) (30) Priority Data: 09/311,482 13 May 1999 (13.05.99) US 09/384,706 27 August 1999 (27.08.99) US (71) Applicant: ZYMOGENETICS, INC. [US/US]; 1201 Eastlake Avenue East, Seattle, WA 98102 (US). (72) Inventors: PIDDINGTON, Christopher, S.; 2838 Irongate Place, Thousand Oaks, CA 91362 (US). PETRIE, Charles, R.; 18459 N.E. 196th Place, Woodinville, WA 98072 (US). SHOEMAKER, Kimberly, E.; 4634 Somerset Avenue S.E., Bellevue, WA 98006 (US). BISHOP, Paul, D.; 28425 S.E. 8th Street, Fall City, WA 98024 (US). (74) Agent: JONES, Phillip, B., C.; ZymoGenetics, Inc., 1201 Eastlake Avenue East, Seattle, WA 98102 (US).		(81) Designated States: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).  Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(54) Title: ZACE2: A HUMAN METALLOENZYME			
(57) Abstract <p>Angiotensin-converting enzyme is a zinc metallopeptidase that plays roles in blood pressure regulation and fertility. The catalytic activities of angiotensin converting enzymes include the production of the potent vasopressor angiotensin II from angiotensin I, and the inactivation of the vasodilatory peptide bradykinin. Zace2 is a new form of human zinc metallopeptidase, characterized as an angiotensin converting enzyme paralog that is expressed primarily in tissues of the digestive system. Two murine orthologs are also described.</p>			

BEST AVAILABLE COPY

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## ZACE2: A HUMAN METALLOENZYME

5

### TECHNICAL FIELD

The present invention relates generally to a new protein expressed by human cells. In particular, the present invention relates to a novel gene that encodes a metalloenzyme, designated as "Zace2," and to nucleic acid molecules encoding Zace2 polypeptides.

### BACKGROUND OF THE INVENTION

Angiotensin-converting enzyme (ACE; peptidyl dipeptidase A; kininase II (EC 3.4.15.1)) is a zinc metallopeptidase that plays roles in blood pressure regulation and fertility. ACE is rather nonspecific and cleaves dipeptides from a broad range of substrates. In general, ACE cleaves a C-terminal dipeptide "A-B" from a polypeptide when A is not a proline residue, and B is neither an aspartate nor a glutamate residue. For example, ACE cleaves a single C-terminal dipeptide from angiotensin I to produce the potent vasopressor angiotensin II, and ACE cleaves the C-terminal dipeptide from [des-Asp<sup>1</sup>]angiotensin I to produce angiotensin III. The enzyme also inactivates the vasodilatory peptide bradykinin by sequential removal of two C-terminal dipeptides. For a general review of angiotensin-converting enzyme, see Corvol *et al.*, *Meth. Enzymol.* 246:283 (1995), Corvol *et al.*, *J. Hypertension* 13(Suppl. 3):S3 (1995), Jackson and Garrison, "Renin and Angiotensin," in *Goodman and Gilman's The Pharmaceutical Basis of Therapeutics*, 9<sup>th</sup> Edition, Molinoff and Ruddon (eds.), pages 733-758 (McGraw-Hill 1996), Matsusaka and Ichikawa, *Annu. Rev. Physiol.* 59:395 (1997), and Zimmerman and Dunham, *Annu. Rev. Pharmacol. Toxicol.* 37:53 (1997).

ACE is a cleavable ectoprotein anchored to the plasma membrane through a transmembrane domain. The majority of the membrane-bound form is extracellularly exposed, and this extracellular domain includes at least one active site. A soluble form of ACE circulates in plasma (see, for example, Hooper and Turner, *Biochem. Soc. Trans.*, 17:660 (1989)).

Two ACE isoforms have been identified in mammalian tissues. The predominant form is referred to as "somatic" ACE, which has a molecular weight of about 150 kD to about 180 kD, and is predominantly found at the surface of vascular endothelial cells, epithelial cells, and neuroepithelial cells. The other isoform is

referred to as "germinal" ACE or testis ACE (tACE), which has a molecular weight of about 90 kD to about 110 kD, and is expressed in post-meiotic cells and sperm. Human somatic ACE has two homologous domains, each comprising a catalytic site and a Zn<sup>2+</sup>-binding region, while human testis ACE contains one catalytic cite.

5           Inhibitors of angiotensin-converting enzyme are used for the treatment of hypertension of various conditions, including left ventricular systolic dysfunction, progressive renal impairment, scleroderma renal crisis, congestive heart failure due to systolic dysfunction, and treatment of atherosclerosis (see, for example, Brown and Vaughan, *Circulation* 97:1411 (1998); Mancini, *Am. J. Med.* 105:40S (1998); Parmley, 10 *Am. J. Med.* 105:27S (1998)). There are at least nine ACE inhibitors approved for use in the United States.

ACE inhibitors can be classified into at least three groups: (1) sulfhydryl-containing inhibitors structurally related to captopril (e.g., fentiapril, pivalopril, zofenopril, alacepril), (2) dicarboxyl-containing inhibitors structurally 15 related to enalapril (e.g., lisinopril, benazepril, quinapril, moexipril, ramipril, spirapril, perindopril, indolapril, pentopril, indalapril, cilazapril), and (3) phosphorus-containing inhibitors structurally related to fosinopril. New classes of ACE inhibitors are sought that will inhibit ACE and other zinc metalloproteases. Moreover, new types of ACE inhibitors are also sought that will selectively inhibit ACE hydrolysis of *N*-acetyl-seryl- 20 aspartyl-lysyl-prolyl (AcSDKP), a regulatory factor in hematopoiesis, without effect on angiotensin I or bradykinin metabolism.

Thus, a continuing need exists for the characterization of new forms of zinc metallopeptidases, and the use of the enzymes to identify therapeutically useful compounds.

25

## BRIEF SUMMARY OF THE INVENTION

The present invention provides a novel metallopeptidase, designated "Zace2." The present invention also provides Zace2 polypeptides and Zace2 fusion proteins, as well as nucleic acid molecules encoding such polypeptides and proteins, 30 and methods for using these nucleic acid molecules and amino acid sequences.

## DETAILED DESCRIPTION OF THE INVENTION

### 1. Overview

A nucleic acid molecule containing a sequence that encodes the human 35 *Zace2* gene has the nucleotide sequence of SEQ ID NO:1. This particular human *Zace2* gene encodes a polypeptide of 805 amino acids (SEQ ID NO:2). Features of *Zace2*

include a putative signal sequence comprising Met<sup>1</sup> through Gln<sup>18</sup> of SEQ ID NO:2, and a transmembrane domain comprising Val<sup>739</sup> through Ile<sup>761</sup> of SEQ ID NO:2. Zace2 also includes a zinc-binding motif that is present in many zinc metalloproteases, and that has the sequence: His-Glu-x-x-His, where "x" is any amino acid (amino acid residues 374 to 378 of SEQ ID NO:2). In this motif, the two histidine residues provide two of the three zinc coordinating ligands, and the glutamate residue is the base donor in the catalytic reaction. The putative third zinc coordinating ligand of Zace2 is Glu<sup>402</sup>. An expanded zinc binding region signature of zinc metalloproteases has the following sequence: [GSTALIVN]-x-x-H-E-[LIVMFYW]-{DEHRKP}-H-x-[LIVMFYWGSPQ], where "x" is any amino acid residue, acceptable amino acid residues are listed between square brackets, and unacceptable amino acid residues are listed between braces (PROSITE sequence No. PS00142 of Release 15.0; Bairoch *et al.*, *Nucleic Acids Res.* 24:217 (1997)). This signature resides within the Zace2 polypeptide at amino acid residues 371 to 380 of SEQ ID NO:2. Zace2 appears to have one catalytic domain, which is similar to the testicular form of ACE.

Two murine Zace2 sequences were also isolated. One of the forms, designated as "mZace-5," has the amino acid sequence of SEQ ID NO:6, and an illustrative nucleotide sequence that encodes the amino acid sequence is provided by SEQ ID NO:5. The amino acid sequence of the second murine form, "mZace2-10," is provided by SEQ ID NO:9, while an exemplary nucleotide sequence that encodes the mZace2-10 polypeptide is included as SEQ ID NO:8. Amino acid sequence analyses revealed that both murine forms share the structural features of the human Zace2 enzyme, as described above.

Northern analysis demonstrated the human Zace2 gene is predominantly expressed by testicular tissue, while there is less expression in kidney, thyroid, small intestine, colon, heart, and potentially, adrenal tissues. In contrast, little or no expression was observed in tissue samples from spleen, thymus, prostate, and ovary. These observations show that Zace2 sequences can be used to differentiate between various tissues.

As described herein, the present invention provides isolated polypeptides comprising an amino acid sequence that is at least 70%, at least 80%, or at least 90% identical to a reference amino acid sequence selected from the group consisting of: (a) amino acid residues 19 to 805 of a sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:6, and SEQ ID NO:9, (b) amino acid residues 19 to 738 of a sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:6, and SEQ ID NO:9, (c) amino acid residues 19 to 708 of a sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:6, and SEQ ID NO:9, (d) amino acid residues

a sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:6, and SEQ ID NO:9, (e) amino acid residues 133 to 542 of a sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:6, and SEQ ID NO:9, (f) amino acid residues 344 to 542 of a sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:6, and SEQ ID NO:9, and (g) amino acid residues 371 to 402 of SEQ ID NO:2 or SEQ ID NO:6, wherein the isolated polypeptide either (a) specifically binds with an antibody that specifically binds with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:6, and SEQ ID NO:9, or (b) exhibits dipeptidyl carboxypeptidase activity. Such a polypeptide can be a metallopeptidase. An illustrative polypeptide is a polypeptide that comprises an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:6, and SEQ ID NO:9. Examples of polypeptides that have amino acid sequences at least 80% identical to amino acid residues 19 to 805 of SEQ ID NO:2 include amino acid residues 19 to 805 of either SEQ ID NO:6 or SEQ ID NO:9.

Additional exemplary polypeptides include polypeptides that comprise an amino acid sequence comprising the motif "[GSTALIVN]-x-x-H-E-[LIVMFYW]-{DEHRKP}-H-x-[LIVMFYWGSPQ]," where "x" is any amino acid residue, acceptable amino acid residues are listed between square brackets, and unacceptable amino acid residues are listed between braces. For example, an illustrative polypeptide comprises amino acid residues 371 to 380 of SEQ ID NO:2.

The present invention also provides isolated polypeptides comprising an extracellular domain, wherein the extracellular domain comprises amino acid residues 19 to 738 of the amino acid sequence of SEQ ID NO:2. Such polypeptides may further comprise a transmembrane domain that resides in a carboxyl-terminal position relative to the extracellular domain, wherein the transmembrane domain comprises amino acid residues 739 to 761 of SEQ ID NO:2. These polypeptides may also comprise an intracellular domain that resides in a carboxyl-terminal position relative to the transmembrane domain, wherein the intracellular domain comprises amino acid residues 762 to 805 of SEQ ID NO:2, and optionally, a signal secretory sequence that resides in an amino-terminal position relative to the extracellular domain, wherein the signal secretory sequence comprises amino acid residues 1 to 18 of the amino acid sequence of SEQ ID NO:2.

The present invention also includes variant Zace2 polypeptides, wherein the amino acid sequence of the variant polypeptide shares an identity with the amino acid sequence of SEQ ID NO:2 selected from the group consisting of at least 70% identity, at least 80% identity, at least 90% identity, at least 95% identity, or greater than 95% identity, and wherein any difference between the amino acid sequence of the

variant polypeptide and the amino acid sequence of SEQ ID NO:2 is due to one or more conservative amino acid substitutions. In addition, the present invention contemplates isolated polypeptides, consisting of an amino acid sequence selected from the group consisting of: (a) amino acid residues 19 to 805 of SEQ ID NO:2, (b) amino acid  
5 residues 19 to 738 of SEQ ID NO:2, (c) amino acid residues 19 to 708 of SEQ ID NO:2, (d) amino acid residues 19 to 613 of SEQ ID NO:2, (e) amino acid residues 133 to 542 of SEQ ID NO:2, (f) amino acid residues 344 to 542 of SEQ ID NO:2, and (g) amino acid residues 371 to 402 of SEQ ID NO:2.

Additional exemplary polypeptides include polypeptides comprising an  
10 amino acid sequence of 15, 20, or 30 contiguous amino acids of an amino acid sequence selected from the group consisting of: amino acid residues 19 to 805 of SEQ ID NO:2, SEQ ID NO:6, or SEQ ID NO:9, amino acid residues 19 to 738 of SEQ ID NO:2, SEQ ID NO:6, or SEQ ID NO:9, amino acid residues 19 to 708 of SEQ ID NO:2, SEQ ID NO:6, or SEQ ID NO:9, amino acid residues 19 to 613 of SEQ ID  
15 NO:2, SEQ ID NO:6, or SEQ ID NO:9, amino acid residues 133 to 542 of SEQ ID NO:2, SEQ ID NO:6, or SEQ ID NO:9, amino acid residues 344 to 542 of SEQ ID NO:2, SEQ ID NO:6, or SEQ ID NO:9, and amino acid residues 371 to 402 of either SEQ ID NO:2 or SEQ ID NO:6. Additional examples of a Zace2 polypeptide include polypeptides consisting of, or comprising, any of the following amino acid sequences:  
20 amino acid residues 19 to 805 of SEQ ID NO:2, SEQ ID NO:6, or SEQ ID NO:9, amino acid residues 19 to 738 of SEQ ID NO:2, SEQ ID NO:6, or SEQ ID NO:9, amino acid residues 19 to 708 of SEQ ID NO:2, SEQ ID NO:6, or SEQ ID NO:9, amino acid residues 19 to 613 of SEQ ID NO:2, SEQ ID NO:6, or SEQ ID NO:9, amino acid residues 133 to 542 of SEQ ID NO:2, SEQ ID NO:6, or SEQ ID NO:9,  
25 amino acid residues 344 to 542 of SEQ ID NO:2, SEQ ID NO:6, or SEQ ID NO:9, and amino acid residues 371 to 402 of either SEQ ID NO:2 or SEQ ID NO:6.

The present invention further provides antibodies and antibody fragments that specifically bind with such polypeptides. Exemplary antibodies include polyclonal antibodies, murine monoclonal antibodies, humanized antibodies derived  
30 from murine monoclonal antibodies, and human monoclonal antibodies. Illustrative antibody fragments include F(ab')<sub>2</sub>, F(ab)<sub>2</sub>, Fab', Fab, Fv, scFv, and minimal recognition units. The present invention further includes compositions comprising a carrier and a protein, peptide, polypeptide, antibody, or anti-idiotypic antibody described herein. For example, the present invention includes pharmaceutical compositions that comprise  
35 such proteins, peptides, polypeptides, antibodies, or anti-idiotypic antibodies, and a pharmaceutically acceptable carrier.

The present invention also provides isolated nucleic acid molecules that encode a *Zace2* polypeptide, wherein the nucleic acid molecule is selected from the group consisting of (a) a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:3, (b) a nucleic acid molecule encoding an amino acid sequence that  
5 comprises amino acid residues 19 to 738 of SEQ ID NO:2, and (c) a nucleic acid molecule that remains hybridized following stringent wash conditions to a nucleic acid molecule comprising the nucleotide sequence of nucleotides 89-2449 of SEQ ID NO:1, or the complement of nucleotides 89-2449 of SEQ ID NO:1.

Illustrative nucleic acid molecules include those in which any difference  
10 between the amino acid sequence encoded by the nucleic acid molecule and the corresponding amino acid sequence of SEQ ID NO:2 is due to a conservative amino acid substitution. The present invention further contemplates isolated nucleic acid molecules that comprise a nucleotide sequence of nucleotides 89-2449 of SEQ ID NO:1, such as a nucleotide sequence of nucleotides 35-2449 of SEQ ID NO:1.

15 The present invention also includes vectors and expression vectors comprising such nucleic acid molecules. Such expression vectors may comprise a transcription promoter, and a transcription terminator, wherein the promoter is operably linked with the nucleic acid molecule, and wherein the nucleic acid molecule is operably linked with the transcription terminator. The present invention further  
20 includes recombinant host cells and recombinant viruses comprising these vectors and expression vectors. Illustrative host cells include bacterial, yeast, avian, fungal, insect, mammalian, and plant cells. Recombinant host cells comprising such expression vectors can be used to produce *Zace2* polypeptides by culturing such recombinant host cells that comprise the expression vector and that produce the *Zace2* protein, and,  
25 optionally, isolating the *Zace2* protein from the cultured recombinant host cells.

In addition, the present invention provides pharmaceutical compositions comprising a pharmaceutically acceptable carrier and at least one of such an expression vector or recombinant virus comprising such expression vectors.

The present invention also contemplates methods for detecting the  
30 presence of *Zace2* RNA in a biological sample, comprising the steps of (a) contacting a *Zace2* nucleic acid probe under hybridizing conditions with either (i) test RNA molecules isolated from the biological sample, or (ii) nucleic acid molecules synthesized from the isolated RNA molecules, wherein the probe has a nucleotide sequence comprising a portion of the nucleotide sequence of SEQ ID NO:1, or its  
35 complement, and (b) detecting the formation of hybrids of the nucleic acid probe and either the test RNA molecules or the synthesized nucleic acid molecules, wherein the presence of the hybrids indicates the presence of *Zace2* RNA in the biological sample.



The present invention further provides methods for detecting the presence of *Zace2* polypeptide in a biological sample, comprising the steps of: (a) contacting the biological sample with an antibody or an antibody fragment that specifically binds with a polypeptide consisting of the amino acid sequence of SEQ ID NO:2, wherein the contacting is performed under conditions that allow the binding of the antibody or antibody fragment to the biological sample, and (b) detecting any of the bound antibody or bound antibody fragment. Such an antibody or antibody fragment may further comprise a detectable label selected from the group consisting of radioisotope, fluorescent label, chemiluminescent label, enzyme label, bioluminescent label, and colloidal gold.

The present invention also provides kits for performing these detection methods. For example, a kit for detection of *Zace2* gene expression may comprise a container that comprises a nucleic acid molecule, wherein the nucleic acid molecule is selected from the group consisting of (a) a nucleic acid molecule comprising the nucleotide sequence of nucleotides 89 to 2449 of SEQ ID NO:1, (b) a nucleic acid molecule comprising the complement of nucleotides 89 to 2449 of the nucleotide sequence of SEQ ID NO:1, (c) a nucleic acid molecule that is a fragment of (a) consisting of at least eight nucleotides, and (d) a nucleic acid molecule that is a fragment of (b) consisting of at least eight nucleotides. Such a kit may also comprise a second container that comprises one or more reagents capable of indicating the presence of the nucleic acid molecule. On the other hand, a kit for detection of *Zace2* protein may comprise a container that comprises an antibody, or an antibody fragment, that specifically binds with a polypeptide consisting of the amino acid sequence of SEQ ID NO:2.

The present invention also contemplates anti-idiotypic antibodies, or anti-idiotypic antibody fragments, that specifically bind an antibody or antibody fragment that specifically binds a polypeptide consisting of the amino acid sequence of SEQ ID NO:2, wherein the anti-idiotypic antibody, or anti-idiotypic antibody fragment, possesses dipeptidyl carboxypeptidase activity.

The present invention also provides isolated nucleic acid molecules comprising a nucleotide sequence that encodes a *Zace2* secretion signal sequence and a nucleotide sequence that encodes a biologically active polypeptide, wherein the *Zace2* secretion signal sequence comprises an amino acid sequence of residues 1 to 18 of SEQ ID NO:2. Illustrative biologically active polypeptides include Factor VIIa, proinsulin, insulin, follicle stimulating hormone, tissue type plasminogen activator, tumor necrosis factor, interleukin, colony stimulating factor, interferon, erythropoietin, and thrombopoietin. Moreover, the present invention provides fusion proteins comprising a

Zace2 secretion signal sequence and a polypeptide, wherein the Zace2 secretion signal sequence comprises an amino acid sequence of residues 1 to 18 of SEQ ID NO:2.

The present invention further includes methods for decreasing inflammation in a subject, comprising the administration of Zace2 or a Zace2 analog, wherein the treatment decreases, in the subject, at least one of vasodilation or serum bradykinin level. For example, such methods can be used to treat a subject when the inflammation is associated with a condition such as inflammatory bowel disease, arthritis, and enterocolitis.

According to the present invention, the Zace2 or Zace2 analog can be administered as a pharmaceutical composition that comprises a Zace2 or Zace2 analog polypeptide and a pharmaceutically acceptable carrier. Suitable pharmaceutical compositions include compositions having a form selected from the group consisting of liquid form, solid form, and aerosol form. Zace2, or an analog, can be a recombinant polypeptide or a polypeptide isolated from a natural source. A suitable Zace2, or Zace2 analog, polypeptide can further comprise a water-soluble polymer, wherein the water-soluble polymer is conjugated to the polypeptide. For example, the water-soluble polymer can be polyethylene glycol.

Zace2 or a Zace2 analog can also be administered as a pharmaceutical composition that comprises a nucleic acid molecule encoding Zace2 or an analog. Moreover, the pharmaceutical composition can comprise at least one of an expression vector that comprises the nucleic acid molecule or a recombinant virus that comprises the expression vector.

The present invention also includes murine Zace2 polypeptides, variant murine Zace2 polypeptides, nucleic acid molecules encoding such polypeptides, expression vectors and recombinant host cells comprising such nucleic acid molecules, and murine Zace2 antibodies and anti-idiotypic antibodies. The present invention also includes compositions and kits comprising murine Zace2 polypeptides and nucleic acid molecules.

The present invention also provides polypeptides comprising at least 25 contiguous amino acid residues of the amino acid sequence of SEQ ID NO:2, SEQ ID NO:6, or SEQ ID NO:9. Illustrative polypeptides include polypeptides comprising at least 50, at least 100, at least 200, or at least 300 amino acid residues of the amino acid sequence of SEQ ID NO:2, SEQ ID NO:6, or SEQ ID NO:9. The present invention further includes nucleic acid molecules that encode such polypeptides.

The present invention further provides fusion proteins comprising a Zace2 moiety and a cell recognition moiety. Suitable cell recognition moieties include receptor ligands, antibodies, and antibody fragments. Other types of fusion proteins

include a Zace2 moiety and an immunoglobulin heavy chain constant region, such as a human F<sub>C</sub> fragment. The present invention further includes isolated nucleic acid molecules that encode such fusion proteins.

5           These and other aspects of the invention will become evident upon reference to the following detailed description. In addition, various references are identified below.

## 2. Definitions

10           In the description that follows, a number of terms are used extensively. The following definitions are provided to facilitate understanding of the invention.

As used herein, "nucleic acid" or "nucleic acid molecule" refers to polynucleotides, such as deoxyribonucleic acid (DNA) or ribonucleic acid (RNA),  
15 oligonucleotides, fragments generated by the polymerase chain reaction (PCR), and fragments generated by any of ligation, scission, endonuclease action, and exonuclease action. Nucleic acid molecules can be composed of monomers that are naturally-occurring nucleotides (such as DNA and RNA), or analogs of naturally-occurring nucleotides (e.g.,  $\alpha$ -enantiomeric forms of naturally-occurring nucleotides), or a  
20 combination of both. Modified nucleotides can have alterations in sugar moieties and/or in pyrimidine or purine base moieties. Sugar modifications include, for example, replacement of one or more hydroxyl groups with halogens, alkyl groups, amines, and azido groups, or sugars can be functionalized as ethers or esters. Moreover, the entire sugar moiety can be replaced with sterically and electronically  
25 similar structures, such as aza-sugars and carbocyclic sugar analogs. Examples of modifications in a base moiety include alkylated purines and pyrimidines, acylated purines or pyrimidines, or other well-known heterocyclic substitutes. Nucleic acid monomers can be linked by phosphodiester bonds or analogs of such linkages. Analogs of phosphodiester linkages include phosphorothioate, phosphorodithioate,  
30 phosphoroselenoate, phosphorodiselenoate, phosphoroanilothioate, phosphoranilidate, phosphoramidate, and the like. The term "nucleic acid molecule" also includes so-called "peptide nucleic acids," which comprise naturally-occurring or modified nucleic acid bases attached to a polyamide backbone. Nucleic acids can be either single stranded or double stranded.

35           The term "complement of a nucleic acid molecule" refers to a nucleic acid molecule having a complementary nucleotide sequence and reverse orientation as

compared to a reference nucleotide sequence. For example, the sequence 5' ATGCACGGG 3' is complementary to 5' CCCGTGCAT 3'.

The term "contig" denotes a nucleic acid molecule that has a contiguous stretch of identical or complementary sequence to another nucleic acid molecule.

5 Contiguous sequences are said to "overlap" a given stretch of a nucleic acid molecule either in their entirety or along a partial stretch of the nucleic acid molecule.

The term "degenerate nucleotide sequence" denotes a sequence of nucleotides that includes one or more degenerate codons as compared to a reference nucleic acid molecule that encodes a polypeptide. Degenerate codons contain different  
10 triplets of nucleotides, but encode the same amino acid residue (*i.e.*, GAU and GAC triplets each encode Asp).

The term "structural gene" refers to a nucleic acid molecule that is transcribed into messenger RNA (mRNA), which is then translated into a sequence of amino acids characteristic of a specific polypeptide.

15 An "isolated nucleic acid molecule" is a nucleic acid molecule that is not integrated in the genomic DNA of an organism. For example, a DNA molecule that encodes a growth factor that has been separated from the genomic DNA of a cell is an isolated DNA molecule. Another example of an isolated nucleic acid molecule is a chemically-synthesized nucleic acid molecule that is not integrated in the genome of an  
20 organism. A nucleic acid molecule that has been isolated from a particular species is smaller than the complete DNA molecule of a chromosome from that species.

A "nucleic acid molecule construct" is a nucleic acid molecule, either single- or double-stranded, that has been modified through human intervention to contain segments of nucleic acid combined and juxtaposed in an arrangement not  
25 existing in nature.

"Linear DNA" denotes non-circular DNA molecules having free 5' and 3' ends. Linear DNA can be prepared from closed circular DNA molecules, such as plasmids, by enzymatic digestion or physical disruption.

"Complementary DNA (cDNA)" is a single-stranded DNA molecule that  
30 is formed from an mRNA template by the enzyme reverse transcriptase. Typically, a primer complementary to portions of mRNA is employed for the initiation of reverse transcription. Those skilled in the art also use the term "cDNA" to refer to a double-stranded DNA molecule consisting of such a single-stranded DNA molecule and its complementary DNA strand. The term "cDNA" also refers to a clone of a cDNA  
35 molecule synthesized from an RNA template.

A "promoter" is a nucleotide sequence that directs the transcription of a structural gene. Typically, a promoter is located in the 5' non-coding region of a gene,

proximal to the transcriptional start site of a structural gene. Sequence elements within promoters that function in the initiation of transcription are often characterized by consensus nucleotide sequences. These promoter elements include RNA polymerase binding sites, TATA sequences, CAAT sequences, differentiation-specific elements (DSEs; McGehee *et al.*, *Mol. Endocrinol.* 7:551 (1993)), cyclic AMP response elements (CREs), serum response elements (SREs; Treisman, *Seminars in Cancer Biol.* 1:47 (1990)), glucocorticoid response elements (GREs), and binding sites for other transcription factors, such as CRE/ATF (O'Reilly *et al.*, *J. Biol. Chem.* 267:19938 (1992)), AP2 (Ye *et al.*, *J. Biol. Chem.* 269:25728 (1994)), SP1, cAMP response element binding protein (CREB; Loeken, *Gene Expr.* 3:253 (1993)) and octamer factors (see, in general, Watson *et al.*, eds., *Molecular Biology of the Gene*, 4th ed. (The Benjamin/Cummings Publishing Company, Inc. 1987), and Lemaigre and Rousseau, *Biochem. J.* 303:1 (1994)). If a promoter is an inducible promoter, then the rate of transcription increases in response to an inducing agent. In contrast, the rate of transcription is not regulated by an inducing agent if the promoter is a constitutive promoter. Repressible promoters are also known.

A "core promoter" contains essential nucleotide sequences for promoter function, including the TATA box and start of transcription. By this definition, a core promoter may or may not have detectable activity in the absence of specific sequences that may enhance the activity or confer tissue specific activity.

A "regulatory element" is a nucleotide sequence that modulates the activity of a core promoter. For example, a regulatory element may contain a nucleotide sequence that binds with cellular factors enabling transcription exclusively or preferentially in particular cells, tissues, or organelles. These types of regulatory elements are normally associated with genes that are expressed in a "cell-specific," "tissue-specific," or "organelle-specific" manner.

An "enhancer" is a type of regulatory element that can increase the efficiency of transcription, regardless of the distance or orientation of the enhancer relative to the start site of transcription.

"Heterologous DNA" refers to a DNA molecule, or a population of DNA molecules, that does not exist naturally within a given host cell. DNA molecules heterologous to a particular host cell may contain DNA derived from the host cell species (*i.e.*, endogenous DNA) so long as that host DNA is combined with non-host DNA (*i.e.*, exogenous DNA). For example, a DNA molecule containing a non-host DNA segment encoding a polypeptide operably linked to a host DNA segment comprising a transcription promoter is considered to be a heterologous DNA molecule. Conversely, a heterologous DNA molecule can comprise an endogenous gene operably

linked with an exogenous promoter. As another illustration, a DNA molecule comprising a gene derived from a wild-type cell is considered to be heterologous DNA if that DNA molecule is introduced into a mutant cell that lacks the wild-type gene.

A "polypeptide" is a polymer of amino acid residues joined by peptide bonds, whether produced naturally or synthetically. Polypeptides of less than about 10 amino acid residues are commonly referred to as "peptides."

A "protein" is a macromolecule comprising one or more polypeptide chains. A protein may also comprise non-peptidic components, such as carbohydrate groups. Carbohydrates and other non-peptidic substituents may be added to a protein by the cell in which the protein is produced, and will vary with the type of cell. Proteins are defined herein in terms of their amino acid backbone structures; substituents such as carbohydrate groups are generally not specified, but may be present nonetheless.

A peptide or polypeptide encoded by a non-host DNA molecule is a "heterologous" peptide or polypeptide.

An "integrated genetic element" is a segment of DNA that has been incorporated into a chromosome of a host cell after that element is introduced into the cell through human manipulation. Within the present invention, integrated genetic elements are most commonly derived from linearized plasmids that are introduced into the cells by electroporation or other techniques. Integrated genetic elements are passed from the original host cell to its progeny.

A "cloning vector" is a nucleic acid molecule, such as a plasmid, cosmid, or bacteriophage, which has the capability of replicating autonomously in a host cell. Cloning vectors typically contain one or a small number of restriction endonuclease recognition sites that allow insertion of a nucleic acid molecule in a determinable fashion without loss of an essential biological function of the vector, as well as nucleotide sequences encoding a marker gene that is suitable for use in the identification and selection of cells transformed with the cloning vector. Marker genes typically include genes that provide tetracycline resistance or ampicillin resistance.

An "expression vector" is a nucleic acid molecule encoding a gene that is expressed in a host cell. Typically, an expression vector comprises a transcription promoter, a gene, and a transcription terminator. Gene expression is usually placed under the control of a promoter, and such a gene is said to be "operably linked to" the promoter. Similarly, a regulatory element and a core promoter are operably linked if the regulatory element modulates the activity of the core promoter.

A "recombinant host" is a cell that contains a heterologous nucleic acid molecule, such as a cloning vector or expression vector. In the present context, an

example of a recombinant host is a cell that produces Zace2 from an expression vector. In contrast, Zace2 can be produced by a cell that is a "natural source" of Zace2, and that lacks an expression vector.

"Integrative transformants" are recombinant host cells, in which  
5 heterologous DNA has become integrated into the genomic DNA of the cells.

A "fusion protein" is a hybrid protein expressed by a nucleic acid molecule comprising nucleotide sequences of at least two genes. For example, a fusion protein can comprise at least part of a Zace2 polypeptide fused with a polypeptide that binds an affinity matrix. Such a fusion protein provides a means to isolate large  
10 quantities of Zace2 using affinity chromatography.

The term "receptor" denotes a cell-associated protein that binds to a bioactive molecule termed a "ligand." This interaction mediates the effect of the ligand on the cell. Receptors can be membrane bound, cytosolic or nuclear; monomeric (e.g., thyroid stimulating hormone receptor, beta-adrenergic receptor) or multimeric (e.g.,  
15 PDGF receptor, growth hormone receptor, IL-3 receptor, GM-CSF receptor, G-CSF receptor, erythropoietin receptor and IL-6 receptor). Membrane-bound receptors are characterized by a multi-domain structure comprising an extracellular ligand-binding domain and an intracellular effector domain that is typically involved in signal transduction. In certain membrane-bound receptors, the extracellular ligand-binding  
20 domain and the intracellular effector domain are located in separate polypeptides that comprise the complete functional receptor.

In general, the binding of ligand to receptor results in a conformational change in the receptor that causes an interaction between the effector domain and other molecule(s) in the cell, which in turn leads to an alteration in the metabolism of the cell.  
25 Metabolic events that are often linked to receptor-ligand interactions include gene transcription, phosphorylation, dephosphorylation, increases in cyclic AMP production, mobilization of cellular calcium, mobilization of membrane lipids, cell adhesion, hydrolysis of inositol lipids and hydrolysis of phospholipids.

The term "secretory signal sequence" denotes a DNA sequence that  
30 encodes a peptide (a "secretory peptide") that, as a component of a larger polypeptide, directs the larger polypeptide through a secretory pathway of a cell in which it is synthesized. The larger polypeptide is commonly cleaved to remove the secretory peptide during transit through the secretory pathway.

An "isolated polypeptide" is a polypeptide that is essentially free from  
35 contaminating cellular components, such as carbohydrate, lipid, or other proteinaceous impurities associated with the polypeptide in nature. Typically, a preparation of isolated polypeptide contains the polypeptide in a highly purified form, i.e., at least about 80%

pure, at least about 90% pure, at least about 95% pure, greater than 95% pure, or greater than 99% pure. One way to show that a particular protein preparation contains an isolated polypeptide is by the appearance of a single band following sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis of the protein preparation and  
5 Coomassie Brilliant Blue staining of the gel. However, the term "isolated" does not exclude the presence of the same polypeptide in alternative physical forms, such as dimers or alternatively glycosylated or derivatized forms.

The terms "amino-terminal" and "carboxyl-terminal" are used herein to denote positions within polypeptides. Where the context allows, these terms are used  
10 with reference to a particular sequence or portion of a polypeptide to denote proximity or relative position. For example, a certain sequence positioned carboxyl-terminal to a reference sequence within a polypeptide is located proximal to the carboxyl terminus of the reference sequence, but is not necessarily at the carboxyl terminus of the complete polypeptide.

15 The term "expression" refers to the biosynthesis of a gene product. For example, in the case of a structural gene, expression involves transcription of the structural gene into mRNA and the translation of mRNA into one or more polypeptides.

The term "splice variant" is used herein to denote alternative forms of RNA transcribed from a gene. Splice variation arises naturally through use of  
20 alternative splicing sites within a transcribed RNA molecule, or less commonly between separately transcribed RNA molecules, and may result in several mRNAs transcribed from the same gene. Splice variants may encode polypeptides having altered amino acid sequence. The term splice variant is also used herein to denote a polypeptide encoded by a splice variant of an mRNA transcribed from a gene.

25 As used herein, the term "immunomodulator" includes cytokines, stem cell growth factors, lymphotoxins, co-stimulatory molecules, hematopoietic factors, and synthetic analogs of these molecules.

The term "complement/anti-complement pair" denotes non-identical moieties that form a non-covalently associated, stable pair under appropriate conditions.  
30 For instance, biotin and avidin (or streptavidin) are prototypical members of a complement/anti-complement pair. Other exemplary complement/anti-complement pairs include receptor/ligand pairs, antibody/antigen (or hapten or epitope) pairs, sense/antisense polynucleotide pairs, and the like. Where subsequent dissociation of the complement/anti-complement pair is desirable, the complement/anti-complement  
35 pair preferably has a binding affinity of less than  $10^9 \text{ M}^{-1}$ .

An "anti-idiotypic antibody" is an antibody that binds with the variable region domain of an immunoglobulin. In the present context, an anti-idiotypic antibody



binds with the variable region of an anti-Zace2 antibody, and thus, an anti-idiotypic antibody mimics an epitope of Zace2.

An "antibody fragment" is a portion of an antibody such as F(ab')<sub>2</sub>, F(ab)<sub>2</sub>, Fab', Fab, and the like. Regardless of structure, an antibody fragment binds with the same antigen that is recognized by the intact antibody. For example, an anti-Zace2 monoclonal antibody fragment binds with an epitope of Zace2.

The term "antibody fragment" also includes a synthetic or a genetically engineered polypeptide that binds to a specific antigen, such as polypeptides consisting of the light chain variable region, "Fv" fragments consisting of the variable regions of the heavy and light chains, recombinant single chain polypeptide molecules in which light and heavy variable regions are connected by a peptide linker ("scFv proteins"), and minimal recognition units consisting of the amino acid residues that mimic the hypervariable region.

A "chimeric antibody" is a recombinant protein that contains the variable domains and complementary determining regions derived from a rodent antibody, while the remainder of the antibody molecule is derived from a human antibody.

"Humanized antibodies" are recombinant proteins in which murine complementarity determining regions of a monoclonal antibody have been transferred from heavy and light variable chains of the murine immunoglobulin into a human variable domain.

As used herein, a "therapeutic agent" is a molecule or atom, which is conjugated to an antibody moiety to produce a conjugate, which is useful for therapy. Examples of therapeutic agents include drugs, toxins, immunomodulators, chelators, boron compounds, photoactive agents or dyes, and radioisotopes.

A "detectable label" is a molecule or atom, which can be conjugated to an antibody moiety to produce a molecule useful for diagnosis. Examples of detectable labels include chelators, photoactive agents, radioisotopes, fluorescent agents, paramagnetic ions, or other marker moieties.

The term "affinity tag" is used herein to denote a polypeptide segment that can be attached to a second polypeptide to provide for purification or detection of the second polypeptide or provide sites for attachment of the second polypeptide to a substrate. In principal, any peptide or protein for which an antibody or other specific binding agent is available can be used as an affinity tag. Affinity tags include a poly-histidine tract, protein A (Nilsson *et al.*, *EMBO J.* 4:1075 (1985); Nilsson *et al.*, *Methods Enzymol.* 198:3 (1991)), glutathione S transferase (Smith and Johnson, *Gene* 67:31 (1988)), Glu-Glu affinity tag (Grussenmeyer *et al.*, *Proc. Natl. Acad. Sci. USA* 82:7952 (1985)), substance P, FLAG peptide (Hopp *et al.*, *Biotechnology* 6:1204

(1988)), streptavidin binding peptide, or other antigenic epitope or binding domain. See, in general, Ford *et al.*, *Protein Expression and Purification* 2:95 (1991). DNA molecules encoding affinity tags are available from commercial suppliers (e.g., Pharmacia Biotech, Piscataway, NJ).

5 A "naked antibody" is an entire antibody, as opposed to an antibody fragment, which is not conjugated with a therapeutic agent. Naked antibodies include both polyclonal and monoclonal antibodies, as well as certain recombinant antibodies, such as chimeric and humanized antibodies.

As used herein, the term "antibody component" includes both an entire  
10 antibody and an antibody fragment.

An "immunoconjugate" is a conjugate of an antibody component with a therapeutic agent or a detectable label.

As used herein, the term "antibody fusion protein" refers to a recombinant molecule that comprises an antibody component and a Zace2 polypeptide  
15 component. An example of an antibody fusion protein is a protein that comprises a Zace2 catalytic domain and either an Fc domain or an antigen-binding region.

A "target polypeptide" or a "target peptide" is an amino acid sequence that comprises at least one epitope, and that is expressed on a target cell, such as a tumor cell, or a cell that carries an infectious agent antigen. T cells recognize peptide  
20 epitopes presented by a major histocompatibility complex molecule to a target polypeptide or target peptide and typically lyse the target cell or recruit other immune cells to the site of the target cell, thereby killing the target cell.

An "antigenic peptide" is a peptide, which will bind a major histocompatibility complex molecule to form an MHC-peptide complex, which is  
25 recognized by a T cell, thereby inducing a cytotoxic lymphocyte response upon presentation to the T cell. Thus, antigenic peptides are capable of binding to an appropriate major histocompatibility complex molecule and inducing a cytotoxic T cells response, such as cell lysis or specific cytokine release against the target cell, which binds or expresses the antigen. The antigenic peptide can be bound in the  
30 context of a class I or class II major histocompatibility complex molecule, on an antigen presenting cell or on a target cell.

In eukaryotes, RNA polymerase II catalyzes the transcription of a structural gene to produce mRNA. A nucleic acid molecule can be designed to contain an RNA polymerase II template in which the RNA transcript has a sequence that is  
35 complementary to that of a specific mRNA. The RNA transcript is termed an "anti-sense RNA" and a nucleic acid molecule that encodes the anti-sense RNA is termed an

“anti-sense gene.” Anti-sense RNA molecules are capable of binding to mRNA molecules, resulting in an inhibition of mRNA translation.

An “anti-sense oligonucleotide specific for *Zace2*” or a “*Zace2* anti-sense oligonucleotide” is an oligonucleotide comprising a nucleotide sequence (a) 5 capable of forming a stable triplex with a portion of the *Zace2* gene, or (b) capable of forming a stable duplex with a portion of an mRNA transcript of the *Zace2* gene.

A “ribozyme” is a nucleic acid molecule that contains a catalytic center. The term includes RNA enzymes, self-splicing RNAs, self-cleaving RNAs, and nucleic acid molecules that perform these catalytic functions. A nucleic acid molecule that 10 encodes a ribozyme is termed a “ribozyme gene.”

An “external guide sequence” is a nucleic acid molecule that directs the endogenous ribozyme, RNase P, to a particular species of intracellular mRNA, resulting in the cleavage of the mRNA by RNase P. A nucleic acid molecule that encodes an external guide sequence is termed an “external guide sequence gene.”

15 The term “human variant *Zace2* gene” refers to nucleic acid molecules that encode a polypeptide comprising an amino acid sequence that is a modification of SEQ ID NO:2. Such variants include naturally-occurring polymorphisms of *Zace2* genes, as well as synthetic genes that contain conservative amino acid substitutions of the amino acid sequence of SEQ ID NO:2. Additional variant forms of *Zace2* genes are 20 nucleic acid molecules that contain insertions or deletions of the nucleotide sequences described herein. A variant *Zace2* gene can be identified by determining whether the gene hybridizes with a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:1, or its complement, under stringent conditions.

Similarly, the term “variant murine *Zace2* gene” refers to nucleic acid 25 molecules that encode a polypeptide having an amino acid sequence that is a modification of SEQ ID NO:6 or 9. A variant murine *Zace2* gene can be identified by determining whether the gene hybridizes with a nucleic acid molecule having the nucleotide sequence of SEQ ID NOs: 5 or 8, or a complement thereof, under stringent conditions.

30 Alternatively, variant *Zace2* genes can be identified by sequence comparison. Two amino acid sequences have “100% amino acid sequence identity” if the amino acid residues of the two amino acid sequences are the same when aligned for maximal correspondence. Similarly, two nucleotide sequences have “100% nucleotide sequence identity” if the nucleotide residues of the two nucleotide sequences are the 35 same when aligned for maximal correspondence. Sequence comparisons can be performed using standard software programs such as those included in the LASERGENE bioinformatics computing suite, which is produced by DNASTAR

(Madison, Wisconsin). Other methods for comparing two nucleotide or amino acid sequences by determining optimal alignment are well-known to those of skill in the art (see, for example, Peruski and Peruski, *The Internet and the New Biology: Tools for Genomic and Molecular Research* (ASM Press, Inc. 1997), Wu *et al.* (eds.),  
5 "Information Superhighway and Computer Databases of Nucleic Acids and Proteins," in *Methods in Gene Biotechnology*, pages 123-151 (CRC Press, Inc. 1997), and Bishop (ed.), *Guide to Human Genome Computing*, 2nd Edition (Academic Press, Inc. 1998)). Particular methods for determining sequence identity are described below.

Regardless of the particular method used to identify a variant *Zace2* gene  
10 or variant *Zace2* polypeptide, a variant gene or polypeptide encoded by a variant gene may be functionally characterized the ability to bind specifically to an anti-*Zace2* antibody or by the dipeptidase (*e.g.*, dipeptidyl carboxypeptidase) activity of the variant *Zace2* polypeptide.

The term "allelic variant" is used herein to denote any of two or more  
15 alternative forms of a gene occupying the same chromosomal locus. Allelic variation arises naturally through mutation, and may result in phenotypic polymorphism within populations. Gene mutations can be silent (no change in the encoded polypeptide) or may encode polypeptides having altered amino acid sequence. The term allelic variant is also used herein to denote a protein encoded by an allelic variant of a gene.

20 The term "ortholog" denotes a polypeptide or protein obtained from one species that is the functional counterpart of a polypeptide or protein from a different species. Sequence differences among orthologs are the result of speciation.

"Paralogs" are distinct but structurally related proteins made by an organism. Paralogs are believed to arise through gene duplication. For example,  $\alpha$ -  
25 globin,  $\beta$ -globin, and myoglobin are paralogs of each other.

The present invention includes functional fragments of *Zace2* genes. Within the context of this invention, a "functional fragment" of a *Zace2* gene refers to a nucleic acid molecule that encodes a portion of a *Zace2* polypeptide, which either has  
30 peptidyl dipeptidase activity or specifically binds with an anti-*Zace2* antibody. For example, a functional fragment of a *Zace2* gene comprises a portion of the nucleotide sequence of SEQ ID NO:1, and encodes a polypeptide that binds with a *Zace2*-specific antibody.

The term "dipeptidyl peptidase" refers to an enzyme that cleaves dipeptides from the amino terminus of a polypeptide, whereas the term "dipeptidyl  
35 carboxypeptidase" refers to an enzyme that cleaves dipeptides from the carboxyl terminus of a polypeptide.

A "metallopeptidase" is a peptide hydrolase, which uses a metal in the catalytic mechanism. Typically, metallopeptidases contain a tightly bound transition metal, such as zinc or iron. Angiotensin-converting enzyme (ACE) is an example of a zinc metallopeptidase. The enzymatic activities of ACE include cleavage of the  
5 carboxyl-terminal dipeptide from angiotensin I to produce angiotensin II, removal of two carboxyl-terminal dipeptides from bradykinin, hydrolysis of *N*-acetyl-Ser-Gly-Lys-Pro at the Gly-Lys bond, cleavage of a carboxyl-terminal tripeptide amide from substance P, and luteinizing hormone releasing hormone, and an amino-terminal tripeptide from luteinizing hormone releasing hormone. Several examples of artificial  
10 ACE substrate are described herein.

Due to the imprecision of standard analytical methods, molecular weights and lengths of polymers are understood to be approximate values. When such a value is expressed as "about" X or "approximately" X, the stated value of X will be understood to be accurate to  $\pm 10\%$ .

15

### 3. *Production of the Zace2 Gene*

Nucleic acid molecules encoding a human *Zace2* gene can be obtained by screening a human cDNA or genomic library using polynucleotide probes based upon SEQ ID NO:1. These techniques are standard and well-established.

20 As an illustration, a nucleic acid molecule that encodes a human *Zace2* gene can be isolated from a cDNA library. In this case, the first step would be to prepare the cDNA library by isolating RNA from a tissue, such testicular tissue, small intestine or colon tissue, using methods well-known to those of skill in the art. In general, RNA isolation techniques must provide a method for breaking cells, a means of  
25 inhibiting RNase-directed degradation of RNA, and a method of separating RNA from DNA, protein, and polysaccharide contaminants. For example, total RNA can be isolated by freezing tissue in liquid nitrogen, grinding the frozen tissue with a mortar and pestle to lyse the cells, extracting the ground tissue with a solution of phenol/chloroform to remove proteins, and separating RNA from the remaining impurities by selective precipitation  
30 with lithium chloride (see, for example, Ausubel *et al.* (eds.), *Short Protocols in Molecular Biology*, 3<sup>rd</sup> Edition, pages 4-1 to 4-6 (John Wiley & Sons 1995) ["Ausubel (1995)"]; Wu *et al.*, *Methods in Gene Biotechnology*, pages 33-41 (CRC Press, Inc. 1997) ["Wu (1997)"]).

Alternatively, total RNA can be isolated by extracting ground tissue with  
35 guanidinium isothiocyanate, extracting with organic solvents, and separating RNA from contaminants using differential centrifugation (see, for example, Chirgwin *et al.*,

*Biochemistry* 18:52 (1979); Ausubel (1995) at pages 4-1 to 4-6; Wu (1997) at pages 33-41).

In order to construct a cDNA library, poly(A)<sup>+</sup> RNA must be isolated from a total RNA preparation. Poly(A)<sup>+</sup> RNA can be isolated from total RNA using the standard technique of oligo(dT)-cellulose chromatography (see, for example, Aviv and Leder, *Proc. Nat'l Acad. Sci. USA* 69:1408 (1972); Ausubel (1995) at pages 4-11 to 4-12).

Double-stranded cDNA molecules are synthesized from poly(A)<sup>+</sup> RNA using techniques well-known to those in the art. (see, for example, Wu (1997) at pages 41-46). Moreover, commercially available kits can be used to synthesize double-stranded cDNA molecules. For example, such kits are available from Life Technologies, Inc. (Gaithersburg, MD), CLONTECH Laboratories, Inc. (Palo Alto, CA), Promega Corporation (Madison, WI) and STRATAGENE (La Jolla, CA).

Various cloning vectors are appropriate for the construction of a cDNA library. For example, a cDNA library can be prepared in a vector derived from bacteriophage, such as a  $\lambda$ gt10 vector. See, for example, Huynh *et al.*, "Constructing and Screening cDNA Libraries in  $\lambda$ gt10 and  $\lambda$ gt11," in *DNA Cloning: A Practical Approach Vol. I*, Glover (ed.), page 49 (IRL Press, 1985); Wu (1997) at pages 47-52.

Alternatively, double-stranded cDNA molecules can be inserted into a plasmid vector, such as a PBLUESCRIPT vector (STRATAGENE; La Jolla, CA), a LAMDA GEM-4 (Promega Corp.) or other commercially available vectors. Suitable cloning vectors also can be obtained from the American Type Culture Collection (Manassas, VA).

To amplify the cloned cDNA molecules, the cDNA library is inserted into a prokaryotic host, using standard techniques. For example, a cDNA library can be introduced into competent *E. coli* DH5 cells, which can be obtained, for example, from Life Technologies, Inc. (Gaithersburg, MD).

A human genomic library can be prepared by means well-known in the art (see, for example, Ausubel (1995) at pages 5-1 to 5-6; Wu (1997) at pages 307-327). Genomic DNA can be isolated by lysing tissue with the detergent Sarkosyl, digesting the lysate with proteinase K, clearing insoluble debris from the lysate by centrifugation, precipitating nucleic acid from the lysate using isopropanol, and purifying resuspended DNA on a cesium chloride density gradient.

DNA fragments that are suitable for the production of a genomic library can be obtained by the random shearing of genomic DNA or by the partial digestion of genomic DNA with restriction endonucleases. Genomic DNA fragments can be inserted into a vector, such as a bacteriophage or cosmid vector, in accordance with conventional

techniques, such as the use of restriction enzyme digestion to provide appropriate termini, the use of alkaline phosphatase treatment to avoid undesirable joining of DNA molecules, and ligation with appropriate ligases. Techniques for such manipulation are well-known in the art (see, for example, Ausubel (1995) at pages 5-1 to 5-6; Wu (1997) at pages 307-  
5 327).

Alternatively, human genomic libraries can be obtained from commercial sources such as Research Genetics (Huntsville, AL) and the American Type Culture Collection (Manassas, VA).

A library containing cDNA or genomic clones can be screened with one or  
10 more polynucleotide probes based upon SEQ ID NO:1, using standard methods (see, for example, Ausubel (1995) at pages 6-1 to 6-11).

Nucleic acid molecules that encode a human *Zace2* gene can also be obtained using the polymerase chain reaction (PCR) with oligonucleotide primers having nucleotide sequences that are based upon the nucleotide sequences of the *Zace2*  
15 gene, as described herein. General methods for screening libraries with PCR are provided by, for example, Yu *et al.*, "Use of the Polymerase Chain Reaction to Screen Phage Libraries," in *Methods in Molecular Biology, Vol. 15: PCR Protocols: Current Methods and Applications*, White (ed.), pages 211-215 (Humana Press, Inc. 1993). Moreover, techniques for using PCR to isolate related genes are described by, for  
20 example, Preston, "Use of Degenerate Oligonucleotide Primers and the Polymerase Chain Reaction to Clone Gene Family Members," in *Methods in Molecular Biology, Vol. 15: PCR Protocols: Current Methods and Applications*, White (ed.), pages 317-337 (Humana Press, Inc. 1993).

Anti-*Zace2* antibodies, produced as described below, can also be used to  
25 isolate DNA sequences that encode human *Zace2* genes from cDNA libraries. For example, the antibodies can be used to screen  $\lambda$ gt11 expression libraries, or the antibodies can be used for immunoscreening following hybrid selection and translation (see, for example, Ausubel (1995) at pages 6-12 to 6-16; Margolis *et al.*, "Screening  $\lambda$  expression libraries with antibody and protein probes," in *DNA Cloning 2: Expression Systems, 2nd Edition*, Glover *et al.* (eds.), pages 1-14 (Oxford University Press 1995)).  
30

As an alternative, a *Zace2* gene can be obtained by synthesizing nucleic acid molecules using mutually priming long oligonucleotides and the nucleotide sequences described herein (see, for example, Ausubel (1995) at pages 8-8 to 8-9). Established techniques using the polymerase chain reaction provide the ability to  
35 synthesize DNA molecules at least two kilobases in length (Adang *et al.*, *Plant Molec. Biol.* 21:1131 (1993), Bambot *et al.*, *PCR Methods and Applications* 2:266 (1993), Dillon *et al.*, "Use of the Polymerase Chain Reaction for the Rapid Construction of

Synthetic Genes," in *Methods in Molecular Biology, Vol. 15: PCR Protocols: Current Methods and Applications*, White (ed.), pages 263-268, (Humana Press, Inc. 1993), and Holowachuk *et al.*, *PCR Methods Appl.* 4:299 (1995)).

The nucleic acid molecules of the present invention can also be synthesized with "gene machines" using protocols such as the phosphoramidite method. If chemically-synthesized double stranded DNA is required for an application such as the synthesis of a gene or a gene fragment, then each complementary strand is made separately. The production of short genes (60 to 80 base pairs) is technically straightforward and can be accomplished by synthesizing the complementary strands and then annealing them. For the production of longer genes (>300 base pairs), however, special strategies may be required, because the coupling efficiency of each cycle during chemical DNA synthesis is seldom 100%. To overcome this problem, synthetic genes (double-stranded) are assembled in modular form from single-stranded fragments that are from 20 to 100 nucleotides in length. For reviews on polynucleotide synthesis, see, for example, Glick and Pasternak, *Molecular Biotechnology, Principles and Applications of Recombinant DNA* (ASM Press 1994), Itakura *et al.*, *Annu. Rev. Biochem.* 53:323 (1984), and Climie *et al.*, *Proc. Nat'l Acad. Sci. USA* 87:633 (1990).

Similar methods can be used to obtain nucleotide sequences, which encode a murine *Zace2* polypeptide.

The sequence of a *Zace2* cDNA or *Zace2* genomic fragment can be determined using standard methods. *Zace2* polynucleotide sequences disclosed herein can also be used as probes or primers to clone 5' non-coding regions of a *Zace2* gene. Promoter elements from a *Zace2* gene can be used to direct the expression of heterologous genes in digestive tract tissues of, for example, transgenic animals or patients treated with gene therapy. The identification of genomic fragments containing a *Zace2* promoter or regulatory element can be achieved using well-established techniques, such as deletion analysis (see, generally, Ausubel (1995)).

Cloning of 5' flanking sequences also facilitates production of *Zace2* proteins by "gene activation," as disclosed in U.S. Patent No. 5,641,670. Briefly, expression of an endogenous *Zace2* gene in a cell is altered by introducing into the *Zace2* locus a DNA construct comprising at least a targeting sequence, a regulatory sequence, an exon, and an unpaired splice donor site. The targeting sequence is a *Zace2* 5' non-coding sequence that permits homologous recombination of the construct with the endogenous *Zace2* locus, whereby the sequences within the construct become operably linked with the endogenous *Zace2* coding sequence. In this way, an endogenous *Zace2* promoter can be replaced or supplemented with other regulatory sequences to provide enhanced, tissue-specific, or otherwise regulated expression.



#### 4. *Production of Zace2 Gene Variants*

The present invention provides a variety of nucleic acid molecules, including DNA and RNA molecules, which encode the Zace2 polypeptides disclosed herein. Those skilled in the art will readily recognize that, in view of the degeneracy of the genetic code, considerable sequence variation is possible among these polynucleotide molecules. SEQ ID NO:3 is a degenerate nucleotide sequence that encompasses all nucleic acid molecules that encode the Zace2 polypeptide of SEQ ID NO:2. Those skilled in the art will recognize that the degenerate sequence of SEQ ID NO:3 also provides all RNA sequences encoding SEQ ID NO:2, by substituting U for T. Thus, the present invention contemplates Zace2 polypeptide-encoding nucleic acid molecules comprising nucleotide 35 to nucleotide 2449 of SEQ ID NO:1, and their RNA equivalents.

Similarly, SEQ ID NO:7 is a degenerate nucleotide sequence that encodes the murine mZace2-5 polypeptide (SEQ ID NO:6). The present invention includes murine Zace2 polypeptide-encoding nucleic acid molecules comprising nucleotide 106 to nucleotide 2520 of SEQ ID NO:5, and their RNA equivalents.

Table 1 sets forth the one-letter codes used within SEQ ID NOs:3 and 7 to denote degenerate nucleotide positions. "Resolutions" are the nucleotides denoted by a code letter. "Complement" indicates the code for the complementary nucleotide(s). For example, the code Y denotes either C or T, and its complement R denotes A or G, A being complementary to T, and G being complementary to C. This information can be used to generate a degenerate sequence encoding murine Zace2-10 polypeptide.

*Table 1*

<b>Nucleotide</b>	<b>Resolution</b>	<b>Complement</b>	<b>Resolution</b>
A	A	T	T
C	C	G	G
G	G	C	C
T	T	A	A
R	A G	Y	C T
Y	C T	R	A G
M	A C	K	G T
K	G T	M	A C
S	C G	S	C G
W	A T	W	A T
H	A C T	D	A G T
B	C G T	V	A C G
V	A C G	B	C G T
D	A G T	H	A C T
N	A C G T	N	A C G T

The degenerate codons used in SEQ ID NOs:3 and 7, encompassing all  
5 possible codons for a given amino acid, are set forth in Table 2.

Table 2

Amino Acid	One Letter Code	Codons	Degenerate Codon
Cys	C	TGC TGT	TGY
Ser	S	AGC AGT TCA TCC TCG TCT	WSN
Thr	T	ACA ACC ACG ACT	ACN
Pro	P	CCA CCC CCG CCT	CCN
Ala	A	GCA GCC GCG GCT	GCN
Gly	G	GGA GGC GGG GGT	GGN
Asn	N	AAC AAT	AAY
Asp	D	GAC GAT	GAY
Glu	E	GAA GAG	GAR
Gln	Q	CAA CAG	CAR
His	H	CAC CAT	CAY
Arg	R	AGA AGG CGA CGC CGG CGT	MGN
Lys	K	AAA AAG	AAR
Met	M	ATG	ATG
Ile	I	ATA ATC ATT	ATH
Leu	L	CTA CTC CTG CTT TTA TTG	YTN
Val	V	GTA GTC GTG GTT	GTN
Phe	F	TTC TTT	TTY
Tyr	Y	TAC TAT	TAY
Trp	W	TGG	TGG
Ter	.	TAA TAG TGA	TRR
Asn Asp	B		RAY
Glu Gln	Z		SAR
Any	X		NNN

One of ordinary skill in the art will appreciate that some ambiguity is introduced in determining a degenerate codon, representative of all possible codons encoding an amino acid. For example, the degenerate codon for serine (WSN) can, in some circumstances, encode arginine (AGR), and the degenerate codon for arginine (MGN) can, in some circumstances, encode serine (AGY). A similar relationship exists between codons encoding phenylalanine and leucine. Thus, some polynucleotides encompassed by the degenerate sequence may encode variant amino acid sequences, but one of ordinary skill in the art can easily identify such variant sequences by reference to the amino acid sequence of SEQ ID NOs:2, 6, or 9. Variant sequences can be readily tested for functionality as described herein.

Different species can exhibit "preferential codon usage." In general, see, Grantham *et al.*, *Nuc. Acids Res.* 8:1893 (1980), Haas *et al.* *Curr. Biol.* 6:315 (1996), Wain-Hobson *et al.*, *Gene* 13:355 (1981), Grosjean and Fiers, *Gene* 18:199 (1982), Holm, *Nuc. Acids Res.* 14:3075 (1986), Ikemura, *J. Mol. Biol.* 158:573 (1982), Sharp and Matassi, *Curr. Opin. Genet. Dev.* 4:851 (1994), Kane, *Curr. Opin. Biotechnol.* 6:494 (1995), and Makrides, *Microbiol. Rev.* 60:512 (1996). As used herein, the term "preferential codon usage" or "preferential codons" is a term of art referring to protein translation codons that are most frequently used in cells of a certain species, thus favoring one or a few representatives of the possible codons encoding each amino acid (See Table 2). For example, the amino acid threonine (Thr) may be encoded by ACA, ACC, ACG, or ACT, but in mammalian cells ACC is the most commonly used codon; in other species, for example, insect cells, yeast, viruses or bacteria, different Thr codons may be preferential. Preferential codons for a particular species can be introduced into the polynucleotides of the present invention by a variety of methods known in the art. Introduction of preferential codon sequences into recombinant DNA can, for example, enhance production of the protein by making protein translation more efficient within a particular cell type or species. Therefore, the degenerate codon sequence disclosed in SEQ ID NOs:3 or 7 serves as a template for optimizing expression of polynucleotides in various cell types and species commonly used in the art and disclosed herein. Sequences containing preferential codons can be tested and optimized for expression in various species, and tested for functionality as disclosed herein.

The present invention further provides variant polypeptides and nucleic acid molecules that represent counterparts from other species (orthologs). These species include, but are not limited to mammalian, avian, amphibian, reptile, fish, insect and other vertebrate and invertebrate species. Of particular interest are Zace2 polypeptides from other mammalian species, including porcine, ovine, bovine, canine,

feline, equine, and other primate polypeptides. Orthologs of human *Zace2* can be cloned using information and compositions provided by the present invention in combination with conventional cloning techniques. For example, a *Zace2* cDNA can be cloned using mRNA obtained from a tissue or cell type that expresses *Zace2* as disclosed herein. Suitable sources of mRNA can be identified by probing northern blots with probes designed from the sequences disclosed herein. A library is then prepared from mRNA of a positive tissue or cell line. Two murine *Zace2* orthologs are described herein, which were isolated as cDNA produced from murine skin RNA.

Those skilled in the art will recognize that the sequence disclosed in SEQ ID NO:1 represents a single allele of human *Zace2*, and that allelic variation and alternative splicing are expected to occur. Allelic variants of this sequence can be cloned by probing cDNA or genomic libraries from different individuals according to standard procedures. Allelic variants of the nucleotide sequence shown in SEQ ID NO:1, including those containing silent mutations and those in which mutations result in amino acid sequence changes, are within the scope of the present invention, as are proteins, which are allelic variants of SEQ ID NO:2. cDNA molecules generated from alternatively spliced mRNAs, which retain the properties of the *Zace2* polypeptide are included within the scope of the present invention, as are polypeptides encoded by such cDNAs and mRNAs. Allelic variants and splice variants of these sequences can be cloned by probing cDNA or genomic libraries from different individuals or tissues according to standard procedures known in the art. The present invention also includes allelic variants and splice variants of the murine *Zace2* forms described herein.

Within certain embodiments of the invention, the isolated nucleic acid molecules can hybridize under stringent conditions to nucleic acid molecules comprising the nucleotide sequence of SEQ ID NO:1, to nucleic acid molecules consisting of the nucleotide sequence of nucleotides 35 to 2449 of SEQ ID NO:1, or to nucleic acid molecules comprising a nucleotide sequence complementary to SEQ ID NO:1 or to nucleotides 35 to 2449 of SEQ ID NO:1. In general, stringent conditions are selected to be about 5°C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength and pH. The  $T_m$  is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe.

A pair of nucleic acid molecules, such as DNA-DNA, RNA-RNA and DNA-RNA, can hybridize if the nucleotide sequences have some degree of complementarity. Hybrids can tolerate mismatched base pairs in the double helix, but the stability of the hybrid is influenced by the degree of mismatch. The  $T_m$  of the mismatched hybrid decreases by 1°C for every 1-1.5% base pair mismatch. Varying

the stringency of the hybridization conditions allows control over the degree of mismatch that will be present in the hybrid. The degree of stringency increases as the hybridization temperature increases and the ionic strength of the hybridization buffer decreases. Stringent hybridization conditions encompass temperatures of about 5-25°C  
5 below the  $T_m$  of the hybrid and a hybridization buffer having up to 1 M  $\text{Na}^+$ . Higher degrees of stringency at lower temperatures can be achieved with the addition of formamide, which reduces the  $T_m$  of the hybrid about 1°C for each 1% formamide in the buffer solution. Generally, such stringent conditions include temperatures of 20-70°C and a hybridization buffer containing up to 6x SSC and 0-50% formamide. A  
10 higher degree of stringency can be achieved at temperatures of from 40-70°C with a hybridization buffer having up to 4x SSC and from 0-50% formamide. Highly stringent conditions typically encompass temperatures of 42-70°C with a hybridization buffer having up to 1x SSC and 0-50% formamide. Different degrees of stringency can be used during hybridization and washing to achieve maximum specific binding to the  
15 target sequence. Typically, the washes following hybridization are performed at increasing degrees of stringency to remove non-hybridized polynucleotide probes from hybridized complexes.

The above conditions are meant to serve as a guide and it is well within the abilities of one skilled in the art to adapt these conditions for use with a particular  
20 polypeptide hybrid. The  $T_m$  for a specific target sequence is the temperature (under defined conditions) at which 50% of the target sequence will hybridize to a perfectly matched probe sequence. Those conditions that influence the  $T_m$  include, the size and base pair content of the polynucleotide probe, the ionic strength of the hybridization solution, and the presence of destabilizing agents in the hybridization solution.  
25 Numerous equations for calculating  $T_m$  are known in the art, and are specific for DNA, RNA and DNA-RNA hybrids and polynucleotide probe sequences of varying length (see, for example, Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Second Edition (Cold Spring Harbor Press 1989); Ausubel *et al.*, (eds.), *Current Protocols in Molecular Biology* (John Wiley and Sons, Inc. 1987); Berger and Kimmel (eds.), *Guide to Molecular Cloning Techniques*, (Academic Press, Inc. 1987); and Wetmur, *Crit. Rev. Biochem. Mol. Biol.* 26:227 (1990)). Sequence analysis software such as OLIGO 6.0 (LSR; Long Lake, MN) and *Primer Premier 4.0* (Premier Biosoft International; Palo Alto, CA), as well as sites on the Internet, are available tools for analyzing a given  
30 sequence and calculating  $T_m$  based on user defined criteria. Such programs can also analyze a given sequence under defined conditions and identify suitable probe sequences. Typically, hybridization of longer polynucleotide sequences, >50 base pairs, is performed at temperatures of about 20-25°C below the calculated  $T_m$ . For

smaller probes, <50 base pairs, hybridization is typically carried out at the  $T_m$  or 5-10°C below. This allows for the maximum rate of hybridization for DNA-DNA and DNA-RNA hybrids.

The length of the polynucleotide sequence influences the rate and stability of hybrid formation. Smaller probe sequences, <50 base pairs, reach equilibrium with complementary sequences rapidly, but may form less stable hybrids. Incubation times of anywhere from minutes to hours can be used to achieve hybrid formation. Longer probe sequences come to equilibrium more slowly, but form more stable complexes even at lower temperatures. Incubations are allowed to proceed overnight or longer. Generally, incubations are carried out for a period equal to three times the calculated Cot time. Cot time, the time it takes for the polynucleotide sequences to reassociate, can be calculated for a particular sequence by methods known in the art.

The base pair composition of polynucleotide sequence will effect the thermal stability of the hybrid complex, thereby influencing the choice of hybridization temperature and the ionic strength of the hybridization buffer. A-T pairs are less stable than G-C pairs in aqueous solutions containing sodium chloride. Therefore, the higher the G-C content, the more stable the hybrid. Even distribution of G and C residues within the sequence also contribute positively to hybrid stability. In addition, the base pair composition can be manipulated to alter the  $T_m$  of a given sequence. For example, 5-methyldeoxycytidine can be substituted for deoxycytidine and 5-bromodeoxuridine can be substituted for thymidine to increase the  $T_m$ , whereas 7-deaza-2'-deoxyguanosine can be substituted for guanosine to reduce dependence on  $T_m$ .

The ionic concentration of the hybridization buffer also affects the stability of the hybrid. Hybridization buffers generally contain blocking agents such as Denhardt's solution (Sigma Chemical Co., St. Louis, Mo.), denatured salmon sperm DNA, tRNA, milk powders (BLOTTO), heparin or SDS, and a  $Na^+$  source, such as SSC (1x SSC: 0.15 M sodium chloride, 15 mM sodium citrate) or SSPE (1x SSPE: 1.8 M NaCl, 10 mM  $NaH_2PO_4$ , 1 mM EDTA, pH 7.7). Typically, hybridization buffers contain from between 10 mM - 1 M  $Na^+$ . The addition of destabilizing or denaturing agents such as formamide, tetralkylammonium salts, guanidinium cations or thiocyanate cations to the hybridization solution will alter the  $T_m$  of a hybrid. Typically, formamide is used at a concentration of up to 50% to allow incubations to be carried out at more convenient and lower temperatures. Formamide also acts to reduce non-specific background when using RNA probes.

As an illustration, a nucleic acid molecule encoding a variant Zace2 polypeptide can be hybridized with a nucleic acid molecule having the nucleotide

sequence of SEQ ID NO:1 (or its complement) at 42°C overnight in a solution comprising 50% formamide, 5x SSC, 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution (100x Denhardt's solution: 2% (w/v) Ficoll 400, 2% (w/v) polyvinylpyrrolidone, and 2% (w/v) bovine serum albumin), 10% dextran sulfate, and 20 µg/ml denatured, sheared salmon sperm DNA. One of skill in the art can devise variations of these hybridization conditions. For example, the hybridization mixture can be incubated at a higher temperature, such as about 65°C, in a solution that does not contain formamide. Moreover, premixed hybridization solutions are available (e.g., EXPRESSHYB Hybridization Solution from CLONTECH Laboratories, Inc.), and hybridization can be performed according to the manufacturer's instructions.

Following hybridization, the nucleic acid molecules can be washed to remove non-hybridized nucleic acid molecules under stringent conditions, or under highly stringent conditions. Typical stringent washing conditions include washing in a solution of 0.5x - 2x SSC with 0.1% sodium dodecyl sulfate (SDS) at 55 - 65°C. That is, nucleic acid molecules encoding a variant Zace2 polypeptide remain hybridized with a nucleic acid molecule consisting of the nucleotide sequence of SEQ ID NO:1 (or its complement) following stringent washing conditions, in which the wash stringency is equivalent to 0.5x - 2x SSC with 0.1% SDS at 55 - 65°C, including 0.5x SSC with 0.1% SDS at 55°C, or 2xSSC with 0.1% SDS at 65°C. One of skill in the art can readily devise equivalent conditions, for example, by substituting SSPE for SSC in the wash solution.

Typical highly stringent washing conditions include washing in a solution of 0.1x - 0.2x SSC with 0.1% sodium dodecyl sulfate (SDS) at 50 - 65°C. In other words, nucleic acid molecules encoding a variant Zace2 polypeptide remain hybridized with a nucleic acid molecule consisting of the nucleotide sequence of SEQ ID NO:1 (or its complement) following highly stringent washing conditions, in which the wash stringency is equivalent to 0.1x - 0.2x SSC with 0.1% SDS at 50 - 65°C, including 0.1x SSC with 0.1% SDS at 50°C, or 0.2xSSC with 0.1% SDS at 65°C.

The present invention also provides isolated Zace2 polypeptides that have a substantially similar sequence identity to the polypeptides of SEQ ID NO:2, or their orthologs. The term "substantially similar sequence identity" is used herein to denote polypeptides having at least 70%, at least 80%, at least 90%, at least 95% or greater than 95% sequence identity to the sequences shown in SEQ ID NO:2, or their orthologs.

The present invention also contemplates Zace2 variant nucleic acid molecules that can be identified using two criteria: a determination of the similarity between the encoded polypeptide with the amino acid sequence of SEQ ID NO:2, and a



hybridization assay, as described above. Such *Zace2* variants include nucleic acid molecules (1) that remain hybridized with a nucleic acid molecule consisting of the nucleotide sequence of SEQ ID NO:1 (or its complement) following stringent washing conditions, in which the wash stringency is equivalent to 0.5x - 2x SSC with 0.1% SDS at 55 - 65°C, and (2) that encode a polypeptide having at least 70%, at least 80%, at least 90%, at least 95% or greater than 95% sequence identity to the amino acid sequence of SEQ ID NO:2. Alternatively, *Zace2* variants can be characterized as nucleic acid molecules (1) that remain hybridized with a nucleic acid molecule consisting of the nucleotide sequence of SEQ ID NO:1 (or its complement) following highly stringent washing conditions, in which the wash stringency is equivalent to 0.1x - 0.2x SSC with 0.1% SDS at 50 - 65°C, and (2) that encode a polypeptide having at least 70%, at least 80%, at least 90%, at least 95% or greater than 95% sequence identity to the amino acid sequence of SEQ ID NO:2.

The present invention further includes murine *Zace2* variant nucleic acid molecules identified by at least one of hybridization analysis and sequence identity determination, with reference to SEQ ID NOs:5, 6, 8, and 9. For example, using the approach discussed above, murine *Zace2* variant nucleic acid molecules can be identified using at least one of three criteria: (1) hybridization with a nucleic acid molecule consisting of the nucleotide sequence of SEQ ID NO:5 or 8 (or its complement) following stringent washing conditions, in which the wash stringency is equivalent to 0.5x - 2x SSC with 0.1% SDS at 55 - 65°C, (2) hybridization with a nucleic acid molecule consisting of the nucleotide sequence of SEQ ID NO:5 or 8 (or its complement) following highly stringent washing conditions, in which the wash stringency is equivalent to 0.1x - 0.2x SSC with 0.1% SDS at 50 - 65°C, and (3) an amino acid percent identity that is at least 70%, at least 80%, at least 90%, at least 95% or greater than 95% sequence identity to the amino acid sequence of SEQ ID NO:6 or 9.

Percent sequence identity is determined by conventional methods. See, for example, Altschul *et al.*, *Bull. Math. Bio.* 48:603 (1986), and Henikoff and Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 (1992). Briefly, two amino acid sequences are aligned to optimize the alignment scores using a gap opening penalty of 10, a gap extension penalty of 1, and the "BLOSUM62" scoring matrix of Henikoff and Henikoff (*ibid.*) as shown in Table 3 (amino acids are indicated by the standard one-letter codes). The percent identity is then calculated as:  $(\text{Total number of identical matches}) / (\text{length of the longer sequence plus the number of gaps introduced into the longer sequence in order to align the two sequences}) \times (100)$ .

Table 3

	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
A	4																			
R	-1	5																		
N	-2	0	6																	
D	-2	-2	1	6																
C	0	-3	-3	-3	9															
Q	-1	1	0	0	-3	5														
E	-1	0	0	2	-4	2	5													
G	0	-2	0	-1	-3	-2	-2	6												
H	-2	0	1	-1	-3	0	0	-2	8											
I	-1	-3	-3	-3	-1	-3	-3	-4	-3	4										
L	-1	-2	-3	-4	-1	-2	-3	-4	-3	2	4									
K	-1	2	0	-1	-3	1	1	-2	-1	-3	-2	5								
M	-1	-1	-2	-3	-1	0	-2	-3	-2	1	2	-1	5							
F	-2	-3	-3	-3	-2	-3	-3	-3	-1	0	0	-3	0	6						
P	-1	-2	-2	-1	-3	-1	-1	-2	-2	-3	-1	-2	-4	7						
S	1	-1	1	0	-1	0	0	0	-1	-2	-2	0	-1	-2	-1	4				
T	0	-1	0	-1	-1	-1	-1	-2	-2	-1	-1	-1	-2	-1	1	5				
W	-3	-3	-4	-4	-2	-2	-3	-2	-2	-3	-2	-3	-1	1	-4	-3	-2	11		
Y	-2	-2	-2	-3	-2	-1	-2	-3	2	-1	-1	-2	-1	3	-3	-2	-2	2	7	
V	0	-3	-3	-3	-1	-2	-2	-3	-3	3	1	-2	1	-1	-2	-2	0	-3	-1	4

5

10

15

20

Those skilled in the art appreciate that there are many established algorithms available to align two amino acid sequences. The "FASTA" similarity search algorithm of Pearson and Lipman is a suitable protein alignment method for examining the level of identity shared by an amino acid sequence disclosed herein and the amino acid sequence of a putative Zace2 variant. The FASTA algorithm is described by Pearson and Lipman, *Proc. Nat'l Acad. Sci. USA* 85:2444 (1988), and by Pearson, *Meth. Enzymol.* 183:63 (1990). Briefly, FASTA first characterizes sequence similarity by identifying regions shared by the query sequence (e.g., SEQ ID NO:2) and a test sequence that have either the highest density of identities (if the ktup variable is 1) or pairs of identities (if ktup=2), without considering conservative amino acid substitutions, insertions, or deletions. The ten regions with the highest density of identities are then rescored by comparing the similarity of all paired amino acids using an amino acid substitution matrix, and the ends of the regions are "trimmed" to include only those residues that contribute to the highest score. If there are several regions with scores greater than the "cutoff" value (calculated by a predetermined formula based upon the length of the sequence and the ktup value), then the trimmed initial regions are examined to determine whether the regions can be joined to form an approximate alignment with gaps. Finally, the highest scoring regions of the two amino acid sequences are aligned using a modification of the Needleman-Wunsch-Sellers algorithm (Needleman and Wunsch, *J. Mol. Biol.* 48:444 (1970); Sellers, *SIAM J. Appl. Math.* 26:787 (1974)), which allows for amino acid insertions and deletions. Illustrative parameters for FASTA analysis are: ktup=1, gap opening penalty=10, gap extension penalty=1, and substitution matrix=BLOSUM62. These parameters can be introduced into a FASTA program by modifying the scoring matrix file ("SMATRIX"), as explained in Appendix 2 of Pearson, *Meth. Enzymol.* 183:63 (1990).

FASTA can also be used to determine the sequence identity of nucleic acid molecules using a ratio as disclosed above. For nucleotide sequence comparisons, the ktup value can range between one to six, preferably from three to six, most preferably three, with other parameters set as described above.

The present invention includes nucleic acid molecules that encode a polypeptide having a conservative amino acid change, compared with the amino acid sequence of SEQ ID NO:2, 6, or 9. That is, variants can be obtained that contain one or more amino acid substitutions of SEQ ID NO:2, 6, or 9, in which an alkyl amino acid is substituted for an alkyl amino acid in a Zace2 amino acid sequence, an aromatic amino acid is substituted for an aromatic amino acid in a Zace2 amino acid sequence, a sulfur-containing amino acid is substituted for a sulfur-containing amino acid in a Zace2 amino acid sequence, a hydroxy-containing amino acid is substituted for a hydroxy-

containing amino acid in a Zace2 amino acid sequence, an acidic amino acid is substituted for an acidic amino acid in a Zace2 amino acid sequence, a basic amino acid is substituted for a basic amino acid in a Zace2 amino acid sequence, or a dibasic monocarboxylic amino acid is substituted for a dibasic monocarboxylic amino acid in a Zace2 amino acid sequence. Among the common amino acids, for example, a "conservative amino acid substitution" is illustrated by a substitution among amino acids within each of the following groups: (1) glycine, alanine, valine, leucine, and isoleucine, (2) phenylalanine, tyrosine, and tryptophan, (3) serine and threonine, (4) aspartate and glutamate, (5) glutamine and asparagine, and (6) lysine, arginine and histidine.

As an illustration, Table 4 lists amino acid substitutions between the human Zace2 and mZace2-5 sequences that are conservative according to the above criteria. One class of human Zace2 variants have at least one amino acid substitution listed in Table 4, while one class of murine Zace-2 mutants have at least one amino acid substitution listed in the table. Comparison of the two murine amino acid sequences identified conservative amino acid substitutions at positions 136 (Arg/Lys) and 493 (His/Arg). Additional conservative amino acid substitutions can be devised by those of skill in the art.

Table 4

Position in Sequence	Human SEQ ID NO:2	Murine SEQ ID NO:6
24	Gln	Asn
59	Val	Ala
66	Gly	Ala
83	Tyr	Phe
93	Val	Ile
107	Val	Ala
209	Val	Ala
212	Val	Ala
228	His	Arg
286	Gly	Ala
335	Asp	Glu
337	Gly	Ala
339	Val	Gly
342	Ala	Val
353	Lys	His
465	Lys	Arg
491	Val	Leu
520	Leu	Ile
577	Lys	Arg
582	Arg	Lys
600	Lys	Arg
641	Tyr	Phe
658	Val	Ile
663	Ile	Val
666	Gly	Leu
679	Ile	Val
694	Ile	Val
698	Thr	Ser

Table 4 (Continued)

Position in Sequence	Human SEQ ID NO:2	Murine SEQ ID NO:6
714	Ala	Val
740	Ser	Thr
745	Val	Ile
751	Gly	Ala
752	Val	Leu
753	Ile	Val
766	Arg	Lys
775	Arg	Lys
798	Thr	Ser
801	Val	Ala

5

The BLOSUM62 table is an amino acid substitution matrix derived from about 2,000 local multiple alignments of protein sequence segments, representing highly conserved regions of more than 500 groups of related proteins (Henikoff and Henikoff, *Proc. Nat'l Acad. Sci. USA* 89:10915 (1992)). Accordingly, the BLOSUM62 substitution frequencies can be used to define conservative amino acid substitutions that may be introduced into the amino acid sequences of the present invention. Although it is possible to design amino acid substitutions based solely upon chemical properties (as discussed above), the language "conservative amino acid substitution" preferably refers to a substitution represented by a BLOSUM62 value of greater than -1. For example, an amino acid substitution is conservative if the substitution is characterized by a BLOSUM62 value of 0, 1, 2, or 3. According to this system, preferred conservative amino acid substitutions are characterized by a BLOSUM62 value of at least 1 (e.g., 1, 2 or 3), while more preferred conservative amino acid substitutions are characterized by a BLOSUM62 value of at least 2 (e.g., 2 or 3).

20

Particular variants of Zace2 are characterized by having at least 70%, at least 80%, at least 90%, at least 95% or greater than 95% sequence identity to the corresponding amino acid sequence (i.e., SEQ ID NOs:2, 6, or 9), wherein the variation in amino acid sequence is due to one or more conservative amino acid substitutions.

Conservative amino acid changes in a *Zace2* gene can be introduced by substituting nucleotides for the nucleotides recited in SEQ ID NOs:1, 5, or 8. Such "conservative amino acid" variants can be obtained, for example, by oligonucleotide-directed mutagenesis, linker-scanning mutagenesis, mutagenesis using the polymerase chain reaction, and the like (see Ausubel (1995) at pages 8-10 to 8-22; and McPherson (ed.), *Directed Mutagenesis: A Practical Approach* (IRL Press 1991)). A variant *Zace2* polypeptide can be identified by the ability to specifically bind anti-*Zace2* antibodies.

The proteins of the present invention can also comprise non-naturally occurring amino acid residues. Non-naturally occurring amino acids include, without limitation, *trans*-3-methylproline, 2,4-methanoproline, *cis*-4-hydroxyproline, *trans*-4-hydroxyproline, *N*-methylglycine, *allo*-threonine, methylthreonine, hydroxyethylcysteine, hydroxyethylhomocysteine, nitroglutamine, homoglutamine, pipecolic acid, thiazolidine carboxylic acid, dehydropyrolidine, 3- and 4-methylproline, 3,3-dimethylproline, *tert*-leucine, norvaline, 2-azaphenylalanine, 3-azaphenylalanine, 4-azaphenylalanine, and 4-fluorophenylalanine. Several methods are known in the art for incorporating non-naturally occurring amino acid residues into proteins. For example, an *in vitro* system can be employed wherein nonsense mutations are suppressed using chemically aminoacylated suppressor tRNAs. Methods for synthesizing amino acids and aminoacylating tRNA are known in the art. Transcription and translation of plasmids containing nonsense mutations is typically carried out in a cell-free system comprising an *E. coli* S30 extract and commercially available enzymes and other reagents. Proteins are purified by chromatography. See, for example, Robertson *et al.*, *J. Am. Chem. Soc.* 113:2722 (1991), Ellman *et al.*, *Methods Enzymol.* 202:301 (1991), Chung *et al.*, *Science* 259:806 (1993), and Chung *et al.*, *Proc. Nat'l Acad. Sci. USA* 90:10145 (1993).

In a second method, translation is carried out in *Xenopus* oocytes by microinjection of mutated mRNA and chemically aminoacylated suppressor tRNAs (Turcatti *et al.*, *J. Biol. Chem.* 271:19991 (1996)). Within a third method, *E. coli* cells are cultured in the absence of a natural amino acid that is to be replaced (*e.g.*, phenylalanine) and in the presence of the desired non-naturally occurring amino acid(s) (*e.g.*, 2-azaphenylalanine, 3-azaphenylalanine, 4-azaphenylalanine, or 4-fluorophenylalanine). The non-naturally occurring amino acid is incorporated into the protein in place of its natural counterpart. See, Koide *et al.*, *Biochem.* 33:7470 (1994). Naturally occurring amino acid residues can be converted to non-naturally occurring species by *in vitro* chemical modification. Chemical modification can be combined with site-directed mutagenesis to further expand the range of substitutions (Wynn and Richards, *Protein Sci.* 2:395 (1993)).

A limited number of non-conservative amino acids, amino acids that are not encoded by the genetic code, non-naturally occurring amino acids, and unnatural amino acids may be substituted for Zace2 amino acid residues.

Essential amino acids in the polypeptides of the present invention can be  
5 identified according to procedures known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (Cunningham and Wells, *Science* 244:1081 (1989), Bass *et al.*, *Proc. Nat'l Acad. Sci. USA* 88:4498 (1991), Coombs and Corey, "Site-Directed Mutagenesis and Protein Engineering," in *Proteins: Analysis and Design*, Angeletti (ed.), pages 259-311 (Academic Press, Inc. 1998)). In the latter technique,  
10 single alanine mutations are introduced at every residue in the molecule, and the resultant mutant molecules are tested for biological activity to identify amino acid residues that are critical to the activity of the molecule. See also, Hilton *et al.*, *J. Biol. Chem.* 271:4699 (1996).

As discussed above, amino acid sequence analysis indicates that the  
15 following amino acids play a role in Zace2 enzymatic activity: His<sup>374</sup>, Glu<sup>375</sup>, His<sup>378</sup>, and Glu<sup>402</sup>. Although sequence analysis can be used to further define the Zace2 active site, domains that play a role in Zace2 activity can also be determined by physical analysis of structure, as determined by such techniques as nuclear magnetic resonance, crystallography, electron diffraction or photoaffinity labeling, in conjunction with  
20 mutation of putative contact site amino acids. See, for example, de Vos *et al.*, *Science* 255:306 (1992), Smith *et al.*, *J. Mol. Biol.* 224:899 (1992), and Wlodaver *et al.*, *FEBS Lett.* 309:59 (1992).

Multiple amino acid substitutions can be made and tested using known methods of mutagenesis and screening, such as those disclosed by Reidhaar-Olson and Sauer (*Science* 241:53 (1988)) or Bowie and Sauer (*Proc. Nat'l Acad. Sci. USA* 86:2152 (1989)). Briefly, these authors disclose methods for simultaneously  
25 randomizing two or more positions in a polypeptide, selecting for functional polypeptide, and then sequencing the mutagenized polypeptides to determine the spectrum of allowable substitutions at each position. Other methods that can be used include phage display (*e.g.*, Lowman *et al.*, *Biochem.* 30:10832 (1991), Ladner *et al.*, U.S. Patent No. 5,223,409, Huse, international publication No. WO 92/06204, and region-directed mutagenesis (Derbyshire *et al.*, *Gene* 46:145 (1986), and Ner *et al.*, *DNA* 7:127, (1988)).

Variants of the disclosed Zace2 nucleotide and polypeptide sequences  
35 can also be generated through DNA shuffling as disclosed by Stemmer, *Nature* 370:389 (1994), Stemmer, *Proc. Nat'l Acad. Sci. USA* 91:10747 (1994), and international publication No. WO 97/20078. Briefly, variant DNA molecules are generated by *in*



*vitro* homologous recombination by random fragmentation of a parent DNA followed by reassembly using PCR, resulting in randomly introduced point mutations. This technique can be modified by using a family of parent DNA molecules, such as allelic variants or DNA molecules from different species, to introduce additional variability into the process. Selection or screening for the desired activity, followed by additional iterations of mutagenesis and assay provides for rapid "evolution" of sequences by selecting for desirable mutations while simultaneously selecting against detrimental changes.

Mutagenesis methods as disclosed herein can be combined with high-throughput, automated screening methods to detect activity of cloned, mutagenized polypeptides in host cells. Mutagenized DNA molecules that encode biologically active polypeptides, or polypeptides that bind with anti-Zace2 antibodies, can be recovered from the host cells and rapidly sequenced using modern equipment. These methods allow the rapid determination of the importance of individual amino acid residues in a polypeptide of interest, and can be applied to polypeptides of unknown structure.

The present invention also includes "functional fragments" of human or murine Zace2 polypeptides and nucleic acid molecules encoding such functional fragments. Routine deletion analyses of nucleic acid molecules can be performed to obtain functional fragments of a nucleic acid molecule that encodes a Zace2 polypeptide. As an illustration, DNA molecules having the nucleotide sequence of SEQ ID NO:1 can be digested with *Bal*31 nuclease to obtain a series of nested deletions. The fragments are then inserted into expression vectors in proper reading frame, and the expressed polypeptides are isolated and tested for the ability to bind anti-Zace2 antibodies. One alternative to exonuclease digestion is to use oligonucleotide-directed mutagenesis to introduce deletions or stop codons to specify production of a desired fragment. Alternatively, particular fragments of a Zace2 gene can be synthesized using the polymerase chain reaction.

This general approach is exemplified by studies on the truncation at either or both termini of interferons have been summarized by Horisberger and Di Marco, *Pharmac. Ther.* 66:507 (1995). Moreover, standard techniques for functional analysis of proteins are described by, for example, Treuter *et al.*, *Molec. Gen. Genet.* 240:113 (1993), Content *et al.*, "Expression and preliminary deletion analysis of the 42 kDa 2-5A synthetase induced by human interferon," in *Biological Interferon Systems, Proceedings of ISIR-TNO Meeting on Interferon Systems*, Cantell (ed.), pages 65-72 (Nijhoff 1987), Herschman, "The EGF Receptor," in *Control of Animal Cell Proliferation, Vol. 1*, Boynton *et al.*, (eds.) pages 169-199 (Academic Press 1985),

Coumailleau *et al.*, *J. Biol. Chem.* 270:29270 (1995); Fukunaga *et al.*, *J. Biol. Chem.* 270:25291 (1995); Yamaguchi *et al.*, *Biochem. Pharmacol.* 50:1295 (1995), and Meisel *et al.*, *Plant Molec. Biol.* 30:1 (1996).

5 An example of a functional fragment of a Zace2 polypeptide is a soluble form of Zace2 that lacks a transmembrane domain. Illustrative Zace2 soluble forms include polypeptides consisting of amino acid residues 1 to 738, 1 to 708, 19 to 738, 19 to 708, 1 to 613, and 19 to 613 of SEQ ID NOs:2, 6, or 9. Additional Zace2 fragments that comprise the active site include amino acid residues 133 to 542, 344 to 542, and 371 to 402 of SEQ ID NOs:2, 6, or 9.

10 The present invention also contemplates functional fragments of a Zace2 gene that have amino acid changes, compared with the amino acid sequence of SEQ ID NOs:2, 6, or 9. A variant Zace2 gene can be identified on the basis of structure by determining the level of identity with nucleotide and amino acid sequence of SEQ ID NOs:2, 6, or 9, as discussed above. An alternative approach to identifying a variant  
15 gene on the basis of structure is to determine whether a nucleic acid molecule encoding a potential variant Zace2 gene can hybridize to a nucleic acid molecule having the nucleotide sequence of SEQ ID NOs:1, 5, or 8, as discussed above.

The present invention also provides polypeptide fragments or peptides comprising an epitope-bearing portion of a Zace2 polypeptide described herein. Such  
20 fragments or peptides may comprise an "immunogenic epitope," which is a part of a protein that elicits an antibody response when the entire protein is used as an immunogen. Immunogenic epitope-bearing peptides can be identified using standard methods (see, for example, Geysen *et al.*, *Proc. Nat'l Acad. Sci. USA* 81:3998 (1983)).

In contrast, polypeptide fragments or peptides may comprise an  
25 "antigenic epitope," which is a region of a protein molecule to which an antibody can specifically bind. Certain epitopes consist of a linear or contiguous stretch of amino acids, and the antigenicity of such an epitope is not disrupted by denaturing agents. It is known in the art that relatively short synthetic peptides that can mimic epitopes of a protein can be used to stimulate the production of antibodies against the protein (see,  
30 for example, Sutcliffe *et al.*, *Science* 219:660 (1983)). Accordingly, antigenic epitope-bearing peptides and polypeptides of the present invention are useful to raise antibodies that bind with the polypeptides described herein.

Antigenic epitope-bearing peptides and polypeptides can contain at least four to ten amino acids, at least ten to fifteen amino acids, or about 15 to about 30  
35 amino acids of SEQ ID NOs:2, 6, or 9. Such epitope-bearing peptides and polypeptides can be produced by fragmenting a Zace2 polypeptide, or by chemical peptide synthesis, as described herein. Moreover, epitopes can be selected by phage display of random

peptide libraries (see, for example, Lane and Stephen, *Curr. Opin. Immunol.* 5:268 (1993), and Cortese *et al.*, *Curr. Opin. Biotechnol.* 7:616 (1996)). Standard methods for identifying epitopes and producing antibodies from small peptides that comprise an epitope are described, for example, by Mole, "Epitope Mapping," in *Methods in Molecular Biology*, Vol. 10, Manson (ed.), pages 105-116 (The Humana Press, Inc. 1992), Price, "Production and Characterization of Synthetic Peptide-Derived Antibodies," in *Monoclonal Antibodies: Production, Engineering, and Clinical Application*, Ritter and Ladyman (eds.), pages 60-84 (Cambridge University Press 1995), and Coligan *et al.* (eds.), *Current Protocols in Immunology*, pages 9.3.1 - 9.3.5 and pages 9.4.1 - 9.4.11 (John Wiley & Sons 1997).

For any Zace2 polypeptide, including variants and fusion proteins, one of ordinary skill in the art can readily generate a fully degenerate polynucleotide sequence encoding that variant using the information set forth in Tables 1 and 2 above. Moreover, those of skill in the art can use standard software to devise Zace2 variants based upon the nucleotide and amino acid sequences described herein. Accordingly, the present invention includes a computer-readable medium encoded with a data structure that provides at least one of the following sequences: SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, and SEQ ID NO:9. Suitable forms of computer-readable media include magnetic media and optically-readable media. Examples of magnetic media include a hard or fixed drive, a random access memory (RAM) chip, a floppy disk, digital linear tape (DLT), a disk cache, and a ZIP disk. Optically readable media are exemplified by compact discs (e.g., CD-read only memory (ROM), CD-rewritable (RW), and CD-recordable), and digital versatile/video discs (DVD) (e.g., DVD-ROM, DVD-RAM, and DVD+RW).

## 5. Production of Zace2 Polypeptides

The polypeptides of the present invention, including full-length polypeptides, functional fragments, and fusion proteins, can be produced in recombinant host cells following conventional techniques. To express a Zace2 gene, a nucleic acid molecule encoding the polypeptide must be operably linked to regulatory sequences that control transcriptional expression in an expression vector and then, introduced into a host cell. In addition to transcriptional regulatory sequences, such as promoters and enhancers, expression vectors can include translational regulatory sequences and a marker gene, which is suitable for selection of cells that carry the expression vector.

Expression vectors that are suitable for production of a foreign protein in eukaryotic cells typically contain (1) prokaryotic DNA elements coding for a bacterial

replication origin and an antibiotic resistance marker to provide for the growth and selection of the expression vector in a bacterial host; (2) eukaryotic DNA elements that control initiation of transcription, such as a promoter; and (3) DNA elements that control the processing of transcripts, such as a transcription termination/polyadenylation sequence. As discussed above, expression vectors can also include nucleotide sequences encoding a secretory sequence that directs the heterologous polypeptide into the secretory pathway of a host cell. For example, a *Zace2* expression vector may comprise a *Zace2* gene and a secretory sequence derived from any secreted gene.

*Zace2* proteins of the present invention may be expressed in mammalian cells. Examples of suitable mammalian host cells include African green monkey kidney cells (Vero; ATCC CRL 1587), human embryonic kidney cells (293-HEK; ATCC CRL 1573), baby hamster kidney cells (BHK-21, BHK-570; ATCC CRL 8544, ATCC CRL 10314), canine kidney cells (MDCK; ATCC CCL 34), Chinese hamster ovary cells (CHO-K1; ATCC CCL61; CHO DG44 (Chasin *et al.*, *Som. Cell. Molec. Genet.* 12:555, 1986)), rat pituitary cells (GH1; ATCC CCL82), HeLa S3 cells (ATCC CCL2.2), rat hepatoma cells (H-4-II-E; ATCC CRL 1548) SV40-transformed monkey kidney cells (COS-1; ATCC CRL 1650) and murine embryonic cells (NIH-3T3; ATCC CRL 1658).

For a mammalian host, the transcriptional and translational regulatory signals may be derived from viral sources, such as adenovirus, bovine papilloma virus, simian virus, or the like, in which the regulatory signals are associated with a particular gene, which has a high level of expression. Suitable transcriptional and translational regulatory sequences also can be obtained from mammalian genes, such as actin, collagen, myosin, and metallothionein genes.

Transcriptional regulatory sequences include a promoter region sufficient to direct the initiation of RNA synthesis. Suitable eukaryotic promoters include the promoter of the mouse *metallothionein I* gene (Hamer *et al.*, *J. Molec. Appl. Genet.* 1:273 (1982)), the *TK* promoter of *Herpes* virus (McKnight, *Cell* 31:355 (1982)), the *SV40* early promoter (Benoist *et al.*, *Nature* 290:304 (1981)), the *Rous* sarcoma virus promoter (Gorman *et al.*, *Proc. Nat'l Acad. Sci. USA* 79:6777 (1982)), the cytomegalovirus promoter (Foelcking *et al.*, *Gene* 45:101 (1980)), and the mouse mammary tumor virus promoter (see, generally, Etcheverry, "Expression of Engineered Proteins in Mammalian Cell Culture," in *Protein Engineering: Principles and Practice*, Cleland *et al.* (eds.), pages 163-181 (John Wiley & Sons, Inc. 1996)).

Alternatively, a prokaryotic promoter, such as the bacteriophage T3 RNA polymerase promoter, can be used to control *Zace2* gene expression in

mammalian cells if the prokaryotic promoter is regulated by a eukaryotic promoter (Zhou *et al.*, *Mol. Cell. Biol.* 10:4529 (1990), and Kaufman *et al.*, *Nucl. Acids Res.* 19:4485 (1991)).

5 An expression vector can be introduced into host cells using a variety of standard techniques including calcium phosphate transfection, liposome-mediated transfection, microprojectile-mediated delivery, electroporation, and the like. Preferably, the transfected cells are selected and propagated to provide recombinant host cells that comprise the expression vector stably integrated in the host cell genome. Techniques for  
10 introducing vectors into eukaryotic cells and techniques for selecting such stable transformants using a dominant selectable marker are described, for example, by Ausubel (1995) and by Murray (ed.), *Gene Transfer and Expression Protocols* (Humana Press 1991).

For example, one suitable selectable marker is a gene that provides resistance to the antibiotic neomycin. In this case, selection is carried out in the  
15 presence of a neomycin-type drug, such as G-418 or the like. Selection systems can also be used to increase the expression level of the gene of interest, a process referred to as "amplification." Amplification is carried out by culturing transfectants in the presence of a low level of the selective agent and then increasing the amount of selective agent to select for cells that produce high levels of the products of the  
20 introduced genes. A suitable amplifiable selectable marker is dihydrofolate reductase, which confers resistance to methotrexate. Other drug resistance genes (*e.g.*, hygromycin resistance, multi-drug resistance, puromycin acetyltransferase) can also be used. Alternatively, markers that introduce an altered phenotype, such as green fluorescent protein, or cell surface proteins such as CD4, CD8, Class I MHC, placental  
25 alkaline phosphatase may be used to sort transfected cells from untransfected cells by such means as FACS sorting or magnetic bead separation technology.

Zace2 polypeptides can also be produced by cultured mammalian cells using a viral delivery system. Exemplary viruses for this purpose include adenovirus, herpesvirus, vaccinia virus and adeno-associated virus (AAV). Adenovirus, a double-  
30 stranded DNA virus, is currently the best studied gene transfer vector for delivery of heterologous nucleic acid (for a review, see Becker *et al.*, *Meth. Cell Biol.* 43:161 (1994), and Douglas and Curiel, *Science & Medicine* 4:44 (1997)). Advantages of the adenovirus system include the accommodation of relatively large DNA inserts, the ability to grow to high-titer, the ability to infect a broad range of mammalian cell types, and flexibility that allows use with a large number of available vectors containing  
35 different promoters.

By deleting portions of the adenovirus genome, larger inserts (up to 7 kb) of heterologous DNA can be accommodated. These inserts can be incorporated into the viral DNA by direct ligation or by homologous recombination with a co-transfected plasmid. An option is to delete the essential *E1* gene from the viral vector, which results in the inability to replicate unless the *E1* gene is provided by the host cell. Adenovirus vector-infected human 293 cells (ATCC Nos. CRL-1573, 45504, 45505), for example, can be grown as adherent cells or in suspension culture at relatively high cell density to produce significant amounts of protein (see Garnier *et al.*, *Cytotechnol.* 15:145 (1994)).

*Zace2* can also be expressed in other higher eukaryotic cells, such as avian, fungal, insect, yeast, or plant cells. The baculovirus system provides an efficient means to introduce cloned *Zace2* genes into insect cells. Suitable expression vectors are based upon the *Autographa californica* multiple nuclear polyhedrosis virus (AcMNPV), and contain well-known promoters such as *Drosophila heat shock protein* (*hsp*) 70 promoter, *Autographa californica* nuclear polyhedrosis virus immediate-early gene promoter (*ie-1*) and the *delayed early* 39K promoter, baculovirus *p10* promoter, and the *Drosophila metallothionein* promoter. A second method of making recombinant baculovirus utilizes a transposon-based system described by Luckow (Luckow, *et al.*, *J. Virol.* 67:4566 (1993)). This system, which utilizes transfer vectors, is sold in the BAC-to-BAC kit (Life Technologies, Rockville, MD). This system utilizes a transfer vector, PFASTBAC (Life Technologies) containing a Tn7 transposon to move the DNA encoding the *Zace2* polypeptide into a baculovirus genome maintained in *E. coli* as a large plasmid called a "bacmid." See, Hill-Perkins and Possee, *J. Gen. Virol.* 71:971 (1990), Bonning, *et al.*, *J. Gen. Virol.* 75:1551 (1994), and Chazenbalk, and Rapoport, *J. Biol. Chem.* 270:1543 (1995). In addition, transfer vectors can include an in-frame fusion with DNA encoding an epitope tag at the C- or N-terminus of the expressed *Zace2* polypeptide, for example, a Glu-Glu epitope tag (Grussenmeyer *et al.*, *Proc. Nat'l Acad. Sci.* 82:7952 (1985)). Using a technique known in the art, a transfer vector containing a *Zace2* gene is transformed into *E. coli*, and screened for bacmids, which contain an interrupted *lacZ* gene indicative of recombinant baculovirus. The bacmid DNA containing the recombinant baculovirus genome is then isolated using common techniques.

The illustrative PFASTBAC vector can be modified to a considerable degree. For example, the polyhedrin promoter can be removed and substituted with the baculovirus basic protein promoter (also known as *Pcor*, p6.9 or MP promoter), which is expressed earlier in the baculovirus infection, and has been shown to be advantageous for expressing secreted proteins (see, for example, Hill-Perkins and

Possee, *J. Gen. Virol.* 71:971 (1990), Bonning, *et al.*, *J. Gen. Virol.* 75:1551 (1994), and Chazenbalk and Rapoport, *J. Biol. Chem.* 270:1543 (1995). In such transfer vector constructs, a short or long version of the basic protein promoter can be used. Moreover, transfer vectors can be constructed, which replace the native Zace2 secretory signal sequences with secretory signal sequences derived from insect proteins. For example, a secretory signal sequence from Ecdysteroid Glucosyltransferase (EGT), honey bee Melittin (Invitrogen Corporation; Carlsbad, CA), or baculovirus gp67 (PharMingen; San Diego, CA) can be used in constructs to replace the native Zace2 secretory signal sequence.

The recombinant virus or bacmid is used to transfect host cells. Suitable insect host cells include cell lines derived from IPLB-Sf-21, a *Spodoptera frugiperda* pupal ovarian cell line, such as Sf9 (ATCC CRL 1711), Sf21AE, and Sf21 (Invitrogen Corporation; San Diego, CA), as well as *Drosophila* Schneider-2 cells, and the HIGH FIVEO cell line (Invitrogen) derived from *Trichoplusia ni* (U.S. Patent No. 5,300,435). Commercially available serum-free media can be used to grow and to maintain the cells. Suitable media are Sf900 II™ (Life Technologies) or ESF 921™ (Expression Systems) for the Sf9 cells; and Ex-cellO405™ (JRH Biosciences, Lenexa, KS) or Express FiveO™ (Life Technologies) for the *T. ni* cells. When recombinant virus is used, the cells are typically grown up from an inoculation density of approximately  $2.5 \times 10^5$  cells to a density of  $1-2 \times 10^6$  cells at which time a recombinant viral stock is added at a multiplicity of infection (MOI) of 0.1 to 10, more typically near 3.

Established techniques for producing recombinant proteins in baculovirus systems are provided by Bailey *et al.*, "Manipulation of Baculovirus Vectors," in *Methods in Molecular Biology, Volume 7: Gene Transfer and Expression Protocols*, Murray (ed.), pages 147-168 (The Humana Press, Inc. 1991), by Patel *et al.*, "The baculovirus expression system," in *DNA Cloning 2: Expression Systems, 2nd Edition*, Glover *et al.* (eds.), pages 205-244 (Oxford University Press 1995), by Ausubel (1995) at pages 16-37 to 16-57, by Richardson (ed.), *Baculovirus Expression Protocols* (The Humana Press, Inc. 1995), and by Lucknow, "Insect Cell Expression Technology," in *Protein Engineering: Principles and Practice*, Cleland *et al.* (eds.), pages 183-218 (John Wiley & Sons, Inc. 1996).

Fungal cells, including yeast cells, can also be used to express the genes described herein. Yeast species of particular interest in this regard include *Saccharomyces cerevisiae*, *Pichia pastoris*, and *Pichia methanolica*. Suitable promoters for expression in yeast include promoters from *GAL1* (galactose), *PGK* (phosphoglycerate kinase), *ADH* (alcohol dehydrogenase), *AOX1* (alcohol oxidase), *HIS4* (histidinol dehydrogenase), and the like. Many yeast cloning vectors have been

designed and are readily available. These vectors include YIp-based vectors, such as YIp5, YRp vectors, such as YRp17, YEpl3 and YCp vectors, such as YCp19. Methods for transforming *S. cerevisiae* cells with exogenous DNA and producing recombinant polypeptides therefrom are disclosed by, for example, Kawasaki, U.S. Patent No. 4,599,311, Kawasaki *et al.*, U.S. Patent No. 4,931,373, Brake, U.S. Patent No. 4,870,008, Welch *et al.*, U.S. Patent No. 5,037,743, and Murray *et al.*, U.S. Patent No. 4,845,075. Transformed cells are selected by phenotype determined by the selectable marker, commonly drug resistance or the ability to grow in the absence of a particular nutrient (e.g., leucine). A suitable vector system for use in *Saccharomyces cerevisiae* is the *POT1* vector system disclosed by Kawasaki *et al.* (U.S. Patent No. 4,931,373), which allows transformed cells to be selected by growth in glucose-containing media. Additional suitable promoters and terminators for use in yeast include those from glycolytic enzyme genes (see, e.g., Kawasaki, U.S. Patent No. 4,599,311, Kingsman *et al.*, U.S. Patent No. 4,615,974, and Bitter, U.S. Patent No. 4,977,092) and alcohol dehydrogenase genes. See also U.S. Patents Nos. 4,990,446, 5,063,154, 5,139,936, and 4,661,454.

Transformation systems for other yeasts, including *Hansenula polymorpha*, *Schizosaccharomyces pombe*, *Kluyveromyces lactis*, *Kluyveromyces fragilis*, *Ustilago maydis*, *Pichia pastoris*, *Pichia methanolica*, *Pichia guilliermondii* and *Candida maltosa* are known in the art. See, for example, Gleeson *et al.*, *J. Gen. Microbiol.* 132:3459 (1986), and Cregg, U.S. Patent No. 4,882,279. *Aspergillus* cells may be utilized according to the methods of McKnight *et al.*, U.S. Patent No. 4,935,349. Methods for transforming *Acremonium chrysogenum* are disclosed by Sumino *et al.*, U.S. Patent No. 5,162,228. Methods for transforming *Neurospora* are disclosed by Lambowitz, U.S. Patent No. 4,486,533.

For example, the use of *Pichia methanolica* as host for the production of recombinant proteins is disclosed by Raymond, U.S. Patent No. 5,716,808, Raymond, U.S. Patent No. 5,736,383, Raymond *et al.*, *Yeast* 14:11-23 (1998), and in international publication Nos. WO 97/17450, WO 97/17451, WO 98/02536, and WO 98/02565. DNA molecules for use in transforming *P. methanolica* will commonly be prepared as double-stranded, circular plasmids, which are preferably linearized prior to transformation. For polypeptide production in *P. methanolica*, it is preferred that the promoter and terminator in the plasmid be that of a *P. methanolica* gene, such as a *P. methanolica* alcohol utilization gene (*AUG1* or *AUG2*). Other useful promoters include those of the dihydroxyacetone synthase (DHAS), formate dehydrogenase (FMD), and catalase (CAT) genes. To facilitate integration of the DNA into the host chromosome, it is preferred to have the entire expression segment of the plasmid flanked at both ends



by host DNA sequences. A suitable selectable marker for use in *Pichia methanolica* is a *P. methanolica ADE2* gene, which encodes phosphoribosyl-5-aminoimidazole carboxylase (AIRC; EC 4.1.1.21), and which allows *ade2* host cells to grow in the absence of adenine. For large-scale, industrial processes where it is desirable to minimize the use of methanol, it is preferred to use host cells in which both methanol utilization genes (*AUG1* and *AUG2*) are deleted. For production of secreted proteins, host cells deficient in vacuolar protease genes (*PEP4* and *PRB1*) are preferred. Electroporation is used to facilitate the introduction of a plasmid containing DNA encoding a polypeptide of interest into *P. methanolica* cells. *P. methanolica* cells can be transformed by electroporation using an exponentially decaying, pulsed electric field having a field strength of from 2.5 to 4.5 kV/cm, preferably about 3.75 kV/cm, and a time constant (t) of from 1 to 40 milliseconds, most preferably about 20 milliseconds.

Expression vectors can also be introduced into plant protoplasts, intact plant tissues, or isolated plant cells. Methods for introducing expression vectors into plant tissue include the direct infection or co-cultivation of plant tissue with *Agrobacterium tumefaciens*, microprojectile-mediated delivery, DNA injection, electroporation, and the like. See, for example, Horsch *et al.*, *Science* 227:1229 (1985), Klein *et al.*, *Biotechnology* 10:268 (1992), and Miki *et al.*, "Procedures for Introducing Foreign DNA into Plants," in *Methods in Plant Molecular Biology and Biotechnology*, Glick *et al.* (eds.), pages 67-88 (CRC Press, 1993).

Alternatively, *Zace2* genes can be expressed in prokaryotic host cells. Suitable promoters that can be used to express *Zace2* polypeptides in a prokaryotic host are well-known to those of skill in the art and include promoters capable of recognizing the T4, T3, Sp6 and T7 polymerases, the P<sub>R</sub> and P<sub>L</sub> promoters of bacteriophage lambda, the *trp*, *recA*, heat shock, *lacUV5*, *tac*, *lpp-lacSpr*, *phoA*, and *lacZ* promoters of *E. coli*, promoters of *B. subtilis*, the promoters of the bacteriophages of *Bacillus*, *Streptomyces* promoters, the *int* promoter of bacteriophage lambda, the *bla* promoter of pBR322, and the CAT promoter of the chloramphenicol acetyl transferase gene. Prokaryotic promoters have been reviewed by Glick, *J. Ind. Microbiol.* 1:277 (1987), Watson *et al.*, *Molecular Biology of the Gene*, 4th Ed. (Benjamin Cummins 1987), and by Ausubel *et al.* (1995).

Illustrative prokaryotic hosts include *E. coli* and *Bacillus subtilis*. Suitable strains of *E. coli* include BL21(DE3), BL21(DE3)pLysS, BL21(DE3)pLysE, DH1, DH4I, DH5, DH5I, DH5IF', DH5IMCR, DH10B, DH10B/p3, DH11S, C600, HB101, JM101, JM105, JM109, JM110, K38, RR1, Y1088, Y1089, CSH18, ER1451, and ER1647 (see, for example, Brown (ed.), *Molecular Biology Labfax* (Academic Press 1991)). Suitable strains of *Bacillus subtilis* include BR151, YB886, MI119,

MI120, and B170 (see, for example, Hardy, "Bacillus Cloning Methods," in *DNA Cloning: A Practical Approach*, Glover (ed.) (IRL Press 1985)).

When expressing a Zace2 polypeptide in bacteria such as *E. coli*, the polypeptide may be retained in the cytoplasm, typically as insoluble granules, or may be directed to the periplasmic space by a bacterial secretion sequence. In the former case, the cells are lysed, and the granules are recovered and denatured using, for example, guanidine isothiocyanate or urea. The denatured polypeptide can then be refolded and dimerized by diluting the denaturant, such as by dialysis against a solution of urea and a combination of reduced and oxidized glutathione, followed by dialysis against a buffered saline solution. In the latter case, the polypeptide can be recovered from the periplasmic space in a soluble and functional form by disrupting the cells (by, for example, sonication or osmotic shock) to release the contents of the periplasmic space and recovering the protein, thereby obviating the need for denaturation and refolding.

Methods for expressing proteins in prokaryotic hosts are well-known to those of skill in the art (see, for example, Williams *et al.*, "Expression of foreign proteins in *E. coli* using plasmid vectors and purification of specific polyclonal antibodies," in *DNA Cloning 2: Expression Systems, 2nd Edition*, Glover *et al.* (eds.), page 15 (Oxford University Press 1995), Ward *et al.*, "Genetic Manipulation and Expression of Antibodies," in *Monoclonal Antibodies: Principles and Applications*, page 137 (Wiley-Liss, Inc. 1995), and Georgiou, "Expression of Proteins in Bacteria," in *Protein Engineering: Principles and Practice*, Cleland *et al.* (eds.), page 101 (John Wiley & Sons, Inc. 1996)).

Standard methods for introducing expression vectors into bacterial, yeast, insect, and plant cells are provided, for example, by Ausubel (1995).

General methods for expressing and recovering foreign protein produced by a mammalian cell system are provided by, for example, Etcheverry, "Expression of Engineered Proteins in Mammalian Cell Culture," in *Protein Engineering: Principles and Practice*, Cleland *et al.* (eds.), pages 163 (Wiley-Liss, Inc. 1996). Standard techniques for recovering protein produced by a bacterial system is provided by, for example, Grisshammer *et al.*, "Purification of over-produced proteins from *E. coli* cells," in *DNA Cloning 2: Expression Systems, 2nd Edition*, Glover *et al.* (eds.), pages 59-92 (Oxford University Press 1995). Established methods for isolating recombinant proteins from a baculovirus system are described by Richardson (ed.), *Baculovirus Expression Protocols* (The Humana Press, Inc. 1995).

As an alternative, polypeptides of the present invention can be synthesized by exclusive solid phase synthesis, partial solid phase methods, fragment

condensation or classical solution synthesis. These synthesis methods are well-known to those of skill in the art (see, for example, Merrifield, *J. Am. Chem. Soc.* 85:2149 (1963), Stewart *et al.*, "Solid Phase Peptide Synthesis" (2nd Edition), (Pierce Chemical Co. 1984), Bayer and Rapp, *Chem. Pept. Prot.* 3:3 (1986), Atherton *et al.*, *Solid Phase Peptide Synthesis: A Practical Approach* (IRL Press 1989), Fields and Colowick, "Solid-Phase Peptide Synthesis," *Methods in Enzymology Volume 289* (Academic Press 1997), and Lloyd-Williams *et al.*, *Chemical Approaches to the Synthesis of Peptides and Proteins* (CRC Press, Inc. 1997)). Variations in total chemical synthesis strategies, such as "native chemical ligation" and "expressed protein ligation" are also standard (see, for example, Dawson *et al.*, *Science* 266:776 (1994), Hackeng *et al.*, *Proc. Nat'l Acad. Sci. USA* 94:7845 (1997), Dawson, *Methods Enzymol.* 287: 34 (1997), Muir *et al.*, *Proc. Nat'l Acad. Sci. USA* 95:6705 (1998), and Severinov and Muir, *J. Biol. Chem.* 273:16205 (1998)).

Peptides and polypeptides of the present invention comprise at least six, at least nine, or at least 15 contiguous amino acid residues of SEQ ID NOs:2, 6, or 9. Within certain embodiments of the invention, the polypeptides comprise 20, 30, 40, 50, 100, or more contiguous residues of SEQ ID NO:2, 6, or 9. Nucleic acid molecules encoding such peptides and polypeptides are useful as polymerase chain reaction primers and probes.

20

## 6. Production of Zace2 Fusion Proteins and Conjugates

Fusion proteins of Zace2 can be used to express Zace2 in a recombinant host, and to isolate the produced Zace2. As described below, particular Zace2 fusion proteins also have uses in diagnosis and therapy.

One type of fusion protein comprises a peptide that guides a Zace2 polypeptide from a recombinant host cell. To direct a Zace2 polypeptide into the secretory pathway of a eukaryotic host cell, a secretory signal sequence (also known as a signal peptide, a leader sequence, prepro sequence or pre sequence) is provided in the Zace2 expression vector. While the secretory signal sequence may be derived from Zace2, a suitable signal sequence may also be derived from another secreted protein or synthesized *de novo*. The secretory signal sequence is operably linked to a Zace2-encoding sequence such that the two sequences are joined in the correct reading frame and positioned to direct the newly synthesized polypeptide into the secretory pathway of the host cell. Secretory signal sequences are commonly positioned 5' to the nucleotide sequence encoding the polypeptide of interest, although certain secretory signal sequences may be positioned elsewhere in the nucleotide sequence of interest

(see, e.g., Welch *et al.*, U.S. Patent No. 5,037,743; Holland *et al.*, U.S. Patent No. 5,143,830).

Although the secretory signal sequence of Zace2 or another protein produced by mammalian cells (e.g., tissue-type plasminogen activator signal sequence, as described, for example, in U.S. Patent No. 5,641,655) is useful for expression of Zace2 in recombinant mammalian hosts, a yeast signal sequence is preferred for expression in yeast cells. Examples of suitable yeast signal sequences are those derived from yeast mating pheromone  $\alpha$ -factor (encoded by the *MFa1* gene), invertase (encoded by the *SUC2* gene), or acid phosphatase (encoded by the *PHO5* gene). See, for example, Romanos *et al.*, "Expression of Cloned Genes in Yeast," in *DNA Cloning 2: A Practical Approach*, 2<sup>nd</sup> Edition, Glover and Hames (eds.), pages 123-167 (Oxford University Press 1995).

In bacterial cells, it is often desirable to express a heterologous protein as a fusion protein to decrease toxicity, increase stability, and to enhance recovery of the expressed protein. For example, Zace2 can be expressed as a fusion protein comprising a glutathione S-transferase polypeptide. Glutathione S-transferase fusion proteins are typically soluble, and easily purifiable from *E. coli* lysates on immobilized glutathione columns. In similar approaches, a Zace2 fusion protein comprising a maltose binding protein polypeptide can be isolated with an amylose resin column, while a fusion protein comprising the C-terminal end of a truncated Protein A gene can be purified using IgG-Sepharose. Established techniques for expressing a heterologous polypeptide as a fusion protein in a bacterial cell are described, for example, by Williams *et al.*, "Expression of Foreign Proteins in *E. coli* Using Plasmid Vectors and Purification of Specific Polyclonal Antibodies," in *DNA Cloning 2: A Practical Approach*, 2<sup>nd</sup> Edition, Glover and Hames (Eds.), pages 15-58 (Oxford University Press 1995). In addition, commercially available expression systems are available. For example, the PINPOINT Xa protein purification system (Promega Corporation; Madison, WI) provides a method for isolating a fusion protein comprising a polypeptide that becomes biotinylated during expression with a resin that comprises avidin.

Peptide tags that are useful for isolating heterologous polypeptides expressed by either prokaryotic or eukaryotic cells include polyHistidine tags (which have an affinity for nickel-chelating resin), *c-myc* tags, calmodulin binding protein (isolated with calmodulin affinity chromatography), substance P, the RYIRS tag (which binds with anti-RYIRS antibodies), the Glu-Glu tag, and the FLAG tag (which binds with anti-FLAG antibodies). See, for example, Luo *et al.*, *Arch. Biochem. Biophys.* 329:215 (1996), Morganti *et al.*, *Biotechnol. Appl. Biochem.* 23:67 (1996), and Zheng

*et al.*, *Gene* 186:55 (1997). Nucleic acid molecules encoding such peptide tags are available, for example, from Sigma-Aldrich Corporation (St. Louis, MO).

The present invention also contemplates that the use of the secretory signal sequence contained in the Zace2 polypeptides of the present invention to direct other polypeptides into the secretory pathway. A signal fusion polypeptide can be made wherein a secretory signal sequence derived from amino acid residues 1 to 18 of SEQ ID NOs:2, 6 or 9 is operably linked to another polypeptide using methods known in the art and disclosed herein. The secretory signal sequence contained in the fusion polypeptides of the present invention is preferably fused amino-terminally to an additional peptide to direct the additional peptide into the secretory pathway. Such constructs have numerous applications known in the art. For example, these novel secretory signal sequence fusion constructs can direct the secretion of an active component of a normally non-secreted protein, such as a receptor. Such fusions may be used in a transgenic animal or in a cultured recombinant host to direct peptides through the secretory pathway. With regard to the latter, exemplary polypeptides include pharmaceutically active molecules such as Factor VIIa, proinsulin, insulin, follicle stimulating hormone, tissue type plasminogen activator, tumor necrosis factor, interleukins (e.g., interleukin-1 (IL-1), IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, and IL-15), colony stimulating factors (e.g., granulocyte-colony stimulating factor (G-CSF) and granulocyte macrophage-colony stimulating factor (GM-CSF)), interferons (e.g., interferons- $\alpha$ , - $\beta$ , - $\gamma$ , - $\omega$ , - $\delta$ , and - $\tau$ ), the stem cell growth factor designated "S1 factor," erythropoietin, and thrombopoietin. The Zace2 secretory signal sequence contained in the fusion polypeptides of the present invention is preferably fused amino-terminally to an additional peptide to direct the additional peptide into the secretory pathway. Fusion proteins comprising a Zace2 secretory signal sequence can be constructed using standard techniques.

Another form of fusion protein comprises a Zace2 polypeptide and an immunoglobulin heavy chain constant region, typically an F<sub>c</sub> fragment, which contains two or three constant region domains and a hinge region but lacks the variable region. As an illustration, Chang *et al.*, U.S. Patent No. 5,723,125, describe a fusion protein comprising a human interferon and a human immunoglobulin Fc fragment. The C-terminal of the interferon is linked to the N-terminal of the Fc fragment by a peptide linker moiety. An example of a peptide linker is a peptide comprising primarily a T cell inert sequence, which is immunologically inert. An exemplary peptide linker has the amino acid sequence: GSGGG SGGGG SGGGG S (SEQ ID NO:4). In this fusion protein, a preferred Fc moiety is a human  $\gamma$ 4 chain, which is stable in solution and has little or no complement activating activity. Accordingly, the present invention

contemplates a Zace2 fusion protein that comprises a Zace2 moiety and a human Fc fragment, wherein the C-terminus of the Zace2 moiety is attached to the N-terminus of the Fc fragment via a peptide linker, such as a peptide consisting of the amino acid sequence of SEQ ID NO:4. The Zace2 moiety can be a Zace2 molecule or a fragment thereof. For example, a fusion protein can comprise a fragment of Zace2 that contains the catalytic domain (e.g., a soluble Zace2 fragment) and an Fc fragment (e.g., a human Fc fragment).

In another variation, a Zace2 fusion protein comprises an IgG sequence, a Zace2 moiety covalently joined to the aminoterminal end of the IgG sequence, and a signal peptide that is covalently joined to the aminoterminal of the Zace2 moiety, wherein the IgG sequence consists of the following elements in the following order: a hinge region, a CH<sub>2</sub> domain, and a CH<sub>3</sub> domain. Accordingly, the IgG sequence lacks a CH<sub>1</sub> domain. The Zace2 moiety displays a Zace2 activity, as described herein, such as the ability to react with a substrate. This general approach to producing fusion proteins that comprise both antibody and nonantibody portions has been described by LaRoche *et al.*, EP 742830 (WO 95/21258).

Fusion proteins comprising a Zace2 moiety and an Fc moiety can be used, for example, as an *in vitro* assay tool. For example, the presence of an Zace2 substrate in a biological sample can be detected using a Zace2-immunoglobulin fusion protein, in which the Zace2 moiety is used to bind the substrate, and a macromolecule, such as Protein A or anti-Fc antibody, is used to bind the fusion protein to a solid support. Such systems can also be used to identify agonists and antagonists that interfere with the binding of Zace2 to its receptor.

Other examples of antibody fusion proteins include polypeptides that comprise an antigen-binding domain and a Zace2 fragment that contains a Zace2 catalytic domain. Such molecules can be used to target particular tissues for the benefit of Zace2 enzymatic activity.

The present invention further provides a variety of other polypeptide fusions. For example, a Zace2 polypeptide (corresponding to the C domain of somatic ACE) can be prepared as a fusion to an N domain of somatic ACE. The native Zace2 signal sequence may also be recombinantly exchanged with the signal sequence of somatic ACE or tACE. Likewise, the transmembrane domain of Zace2 may be recombinantly exchanged with that of somatic ACE or tACE. The catalytic domain of Zace2 can also be recombinantly exchanged for the corresponding region of somatic ACE, tACE, thermolysin or another zinc metalloprotease. Accordingly, part or all of a domain(s) conferring a biological function can be swapped between Zace2 of the present invention with the functionally equivalent domain(s) from another family

member, such as tACE or somatic ACE. Polypeptide fusions can be expressed in recombinant host cells to produce a variety of Zace2 fusion analogs. A Zace2 polypeptide can be fused to two or more moieties or domains, such as an affinity tag for purification and a targeting domain. Polypeptide fusions can also comprise one or  
5 more cleavage sites, particularly between domains. See, for example, Tuan *et al.*, *Connective Tissue Research* 34:1 (1996).

Fusion proteins can be prepared by methods known to those skilled in the art by preparing each component of the fusion protein and chemically conjugating them. Alternatively, a polynucleotide encoding both components of the fusion protein  
10 in the proper reading frame can be generated using known techniques and expressed by the methods described herein. General methods for enzymatic and chemical cleavage of fusion proteins are described, for example, by Ausubel (1995) at pages 16-19 to 16-25.

The present invention also contemplates chemically modified Zace2 compositions, in which a Zace2 polypeptide is linked with a polymer. Preferred Zace2 polypeptides are soluble polypeptides that lack a functional transmembrane domain. Typically, the polymer is water soluble so that the Zace2 conjugate does not precipitate in an aqueous environment, such as a physiological environment. An example of a suitable polymer is one that has been modified to have a single reactive group, such as  
20 an active ester for acylation, or an aldehyde for alkylation. In this way, the degree of polymerization can be controlled. An example of a reactive aldehyde is polyethylene glycol propionaldehyde, or mono-(C1-C10) alkoxy, or aryloxy derivatives thereof (see, for example, Harris, *et al.*, U.S. Patent No. 5,252,714). The polymer may be branched or unbranched. Moreover, a mixture of polymers can be used to produce Zace2  
25 conjugates.

Zace2 conjugates used for therapy can comprise pharmaceutically acceptable water-soluble polymer moieties. Suitable water-soluble polymers include polyethylene glycol (PEG), monomethoxy-PEG, mono-(C1-C10)alkoxy-PEG, aryloxy-PEG, poly-(N-vinyl pyrrolidone)PEG, tresyl monomethoxy PEG, PEG  
30 propionaldehyde, *bis*-succinimidyl carbonate PEG, propylene glycol homopolymers, a polypropylene oxide/ethylene oxide co-polymer, polyoxyethylated polyols (*e.g.*, glycerol), polyvinyl alcohol, dextran, cellulose, or other carbohydrate-based polymers. Suitable PEG may have a molecular weight from about 600 to about 60,000, including, for example, 5,000, 12,000, 20,000 and 25,000. A Zace2 conjugate can also comprise a  
35 mixture of such water-soluble polymers.

One example of a Zace2 conjugate comprises a Zace2 moiety and a polyalkyl oxide moiety attached to the *N*-terminus of the Zace2 moiety. PEG is one

suitable polyalkyl oxide. As an illustration, Zace2 can be modified with PEG, a process known as "PEGylation." PEGylation of Zace2 can be carried out by any of the PEGylation reactions known in the art (see, for example, EP 0 154 316, Delgado *et al.*, *Critical Reviews in Therapeutic Drug Carrier Systems* 9:249 (1992), Duncan and Spreafico, *Clin. Pharmacokinet.* 27:290 (1994), and Francis *et al.*, *Int J Hematol* 68:1 (1998)). For example, PEGylation can be performed by an acylation reaction or by an alkylation reaction with a reactive polyethylene glycol molecule. In an alternative approach, Zace2 conjugates are formed by condensing activated PEG, in which a terminal hydroxy or amino group of PEG has been replaced by an activated linker (see, for example, Karasiewicz *et al.*, U.S. Patent No. 5,382,657).

PEGylation by acylation typically requires reacting an active ester derivative of PEG with a Zace2 polypeptide. An example of an activated PEG ester is PEG esterified to *N*-hydroxysuccinimide. As used herein, the term "acylation" includes the following types of linkages between Zace2 and a water soluble polymer: amide, carbamate, urethane, and the like. Methods for preparing PEGylated Zace2 by acylation will typically comprise the steps of (a) reacting a Zace2 polypeptide with PEG (such as a reactive ester of an aldehyde derivative of PEG) under conditions whereby one or more PEG groups attach to Zace2, and (b) obtaining the reaction product(s). Generally, the optimal reaction conditions for acylation reactions will be determined based upon known parameters and desired results. For example, the larger the ratio of PEG:Zace2, the greater the percentage of polyPEGylated Zace2 product.

The product of PEGylation by acylation is typically a polyPEGylated Zace2 product, wherein the lysine  $\epsilon$ -amino groups are PEGylated via an acyl linking group. An example of a connecting linkage is an amide. Typically, the resulting Zace2 will be at least 95% mono-, di-, or tri-pegylated, although some species with higher degrees of PEGylation may be formed depending upon the reaction conditions. PEGylated species can be separated from unconjugated Zace2 polypeptides using standard purification methods, such as dialysis, ultrafiltration, ion exchange chromatography, affinity chromatography, and the like.

PEGylation by alkylation generally involves reacting a terminal aldehyde derivative of PEG with Zace2 in the presence of a reducing agent. PEG groups are preferably attached to the polypeptide via a  $-\text{CH}_2\text{-NH}$  group.

Derivatization via reductive alkylation to produce a monoPEGylated product takes advantage of the differential reactivity of different types of primary amino groups available for derivatization. Typically, the reaction is performed at a pH that allows one to take advantage of the pKa differences between the  $\epsilon$ -amino groups of the lysine residues and the  $\alpha$ -amino group of the *N*-terminal residue of the protein. By



such selective derivatization, attachment of a water-soluble polymer that contains a reactive group such as an aldehyde, to a protein is controlled. The conjugation with the polymer occurs predominantly at the *N*-terminus of the protein without significant modification of other reactive groups such as the lysine side chain amino groups. The present invention provides a substantially homogenous preparation of Zace2 monopolymer conjugates.

Reductive alkylation to produce a substantially homogenous population of monopolymer Zace2 conjugate molecule can comprise the steps of: (a) reacting a Zace2 polypeptide with a reactive PEG under reductive alkylation conditions at a pH suitable to permit selective modification of the  $\alpha$ -amino group at the amino terminus of the Zace2, and (b) obtaining the reaction product(s). The reducing agent used for reductive alkylation should be stable in aqueous solution and preferably be able to reduce only the Schiff base formed in the initial process of reductive alkylation. Preferred reducing agents include sodium borohydride, sodium cyanoborohydride, dimethylamine borane, trimethylamine borane, and pyridine borane.

For a substantially homogenous population of monopolymer Zace2 conjugates, the reductive alkylation reaction conditions are those which permit the selective attachment of the water soluble polymer moiety to the *N*-terminus of Zace2. Such reaction conditions generally provide for pKa differences between the lysine amino groups and the  $\alpha$ -amino group at the *N*-terminus. The pH also affects the ratio of polymer to protein to be used. In general, if the pH is lower, a larger excess of polymer to protein will be desired because the less reactive the *N*-terminal  $\alpha$ -group, the more polymer is needed to achieve optimal conditions. If the pH is higher, the polymer:Zace2 need not be as large because more reactive groups are available. Typically, the pH will fall within the range of 3 to 9, or 3 to 6.

Another factor to consider is the molecular weight of the water-soluble polymer. Generally, the higher the molecular weight of the polymer, the fewer number of polymer molecules, which may be attached to the protein. For PEGylation reactions, the typical molecular weight is about 2 kDa to about 100 kDa, about 5 kDa to about 50 kDa, or about 12 kDa to about 25 kDa. The molar ratio of water-soluble polymer to Zace2 will generally be in the range of 1:1 to 100:1. Typically, the molar ratio of water-soluble polymer to Zace2 will be 1:1 to 20:1 for polyPEGylation, and 1:1 to 5:1 for monoPEGylation.

General methods for producing conjugates comprising a polypeptide and water-soluble polymer moieties are known in the art. See, for example, Karasiewicz *et al.*, U.S. Patent No. 5,382,657, Greenwald *et al.*, U.S. Patent No. 5,738, 846, Nieforth

*et al., Clin. Pharmacol. Ther.* 59:636 (1996), Monkarsh *et al., Anal. Biochem.* 247:434 (1997)).

The present invention contemplates compositions comprising a peptide or polypeptide described herein. Such compositions can further comprise a carrier.  
5 The carrier can be a conventional organic or inorganic carrier. Examples of carriers include water, buffer solution, alcohol, propylene glycol, macrogol, sesame oil, corn oil, and the like.

## 7. Isolation of Zace2 Polypeptides

10 The polypeptides of the present invention to can be purified to at least about 80% purity, to at least about 90% purity, or to at least about 95% purity, or even greater than 95% purity with respect to contaminating macromolecules, particularly other proteins and nucleic acids, and free of infectious and pyrogenic agents. The polypeptides of the present invention may also be purified to a pharmaceutically pure  
15 state, which is greater than 99.9% pure. In certain preparations, a purified polypeptide is substantially free of other polypeptides, particularly other polypeptides of animal origin.

Fractionation and/or conventional purification methods can be used to obtain preparations of Zace2 purified from natural sources (*e.g.*, testicular tissue),  
20 synthetic Zace2 polypeptides, and recombinant Zace2 polypeptides and fusion Zace2 polypeptides purified from recombinant host cells. In general, ammonium sulfate precipitation and acid or chaotrope extraction may be used for fractionation of samples. Exemplary purification steps may include hydroxyapatite, size exclusion, FPLC and reverse-phase high performance liquid chromatography. Suitable chromatographic  
25 media include derivatized dextrans, agarose, cellulose, polyacrylamide, specialty silicas, and the like. PEI, DEAE, QAE and Q derivatives are preferred. Exemplary chromatographic media include those media derivatized with phenyl, butyl, or octyl groups, such as Phenyl-Sepharose FF (Pharmacia), Toyopearl butyl 650 (Toso Haas, Montgomeryville, PA), Octyl-Sepharose (Pharmacia) and the like; or polyacrylic resins,  
30 such as Amberchrom CG 71 (Toso Haas) and the like. Suitable solid supports include glass beads, silica-based resins, cellulosic resins, agarose beads, cross-linked agarose beads, polystyrene beads, cross-linked polyacrylamide resins and the like that are insoluble under the conditions in which they are to be used. These supports may be modified with reactive groups that allow attachment of proteins by amino groups,  
35 carboxyl groups, sulfhydryl groups, hydroxyl groups and/or carbohydrate moieties.

Examples of coupling chemistries include cyanogen bromide activation, N-hydroxysuccinimide activation, epoxide activation, sulfhydryl activation, hydrazide activation, and carboxyl and amino derivatives for carbodiimide coupling chemistries. These and other solid media are well known and widely used in the art, and are available from commercial suppliers. Selection of a particular method for polypeptide isolation and purification is a matter of routine design and is determined in part by the properties of the chosen support. See, for example, *Affinity Chromatography: Principles & Methods* (Pharmacia LKB Biotechnology 1988), and Doonan, *Protein Purification Protocols* (The Humana Press 1996).

Additional variations in Zace2 isolation and purification can be devised by those of skill in the art. For example, anti-Zace2 antibodies, obtained as described below, can be used to isolate large quantities of protein by immunoaffinity purification.

Moreover, methods for binding enzymes, such as Zace2, to substrates bound to support media are well known in the art. For example, the polypeptides of the present invention can be isolated by exploitation of their homology to somatic ACE and tACE. These enzymes can be purified by affinity chromatography using the ACE inhibitor lisinopril [*N*-[(*S*)-1-carboxy-3-phenylpropyl]-Lys-Pro] as the ligand affixed to a solid support. Improved purification yields can be obtained using a 28 Å, rather than a 14 Å, spacer between ligand and solid support.

The polypeptides of the present invention can also be isolated by exploitation of particular properties. For example, immobilized metal ion adsorption (IMAC) chromatography can be used to purify histidine-rich proteins, including those comprising polyhistidine tags. Briefly, a gel is first charged with divalent metal ions to form a chelate (Sulkowski, *Trends in Biochem.* 3:1 (1985)). Histidine-rich proteins will be adsorbed to this matrix with differing affinities, depending upon the metal ion used, and will be eluted by competitive elution, lowering the pH, or use of strong chelating agents. Other methods of purification include purification of glycosylated proteins by lectin affinity chromatography and ion exchange chromatography (M. Deutscher, (ed.), *Meth. Enzymol.* 182:529 (1990)). Within additional embodiments of the invention, a fusion of the polypeptide of interest and an affinity tag (*e.g.*, maltose-binding protein, an immunoglobulin domain) may be constructed to facilitate purification.

Zace2 polypeptides or fragments thereof may also be prepared through chemical synthesis, as described below. Zace2 polypeptides may be monomers or multimers; glycosylated or non-glycosylated; PEGylated or non-PEGylated; and may or may not include an initial methionine amino acid residue.

## 8. Zace2 Analogs and Zace2 Inhibitors

One general class of Zace2 analogs are variants having an amino acid sequence that is a mutation of the amino acid sequence disclosed herein. Another general class of Zace2 analogs is provided by anti-idiotypic antibodies, and fragments thereof, as described below. Moreover, recombinant antibodies comprising anti-idiotypic variable domains can be used as analogs (see, for example, Monfardini *et al.*, *Proc. Assoc. Am. Physicians* 108:420 (1996)). Since the variable domains of anti-idiotypic Zace2 antibodies mimic Zace2, these domains can provide Zace2 enzymatic activity. Methods of producing anti-idiotypic catalytic antibodies are known to those of skill in the art (see, for example, Joron *et al.*, *Ann. N Y Acad. Sci.* 672:216 (1992), Friboulet *et al.*, *Appl. Biochem. Biotechnol.* 47:229 (1994), and Avelle *et al.*, *Ann. N Y Acad. Sci.* 864:118 (1998)).

Another approach to identifying Zace2 analogs is provided by the use of combinatorial libraries. Methods for constructing and screening phage display and other combinatorial libraries are provided, for example, by Kay *et al.*, *Phage Display of Peptides and Proteins* (Academic Press 1996), Verdine, U.S. Patent No. 5,783,384, Kay, *et al.*, U.S. Patent No. 5,747,334, and Kauffman *et al.*, U.S. Patent No. 5,723,323.

One illustrative *in vitro* use of Zace2 and its analogs is the production of labeled angiotensin II. For example, angiotensin I, radiolabeled at its N-terminus, can be incubated in the presence of Zace2 or an active variant Zace2. The product of the reaction will be radiolabeled angiotensin II. This radiolabeled molecule can be used to study the metabolism of angiotensin II *in vitro*, or to observe the tissue distribution of administered angiotensin II *in vivo*.

The activity of Zace2 molecules of the present invention can be measured using a variety of assays that measure catalytic activity of the enzyme in the presence or absence of zinc, or that measure the effects of chloride or other monoanions on the catalytic activity of Zace2. In addition, the Zace2 polypeptides can be characterized by measuring the zinc content of these polypeptides. Radiolabeled ACE inhibitors are useful for detecting high-affinity binding sites in zinc metalloprotease family members. One or more mutations of putative critical or important residues, in conjunction with known assays of ACE activity, permit analysis of mutational effects on Zace2 structure, enzyme activity, and immunological activity. In addition, both synthetic and natural ACE substrates can be useful in characterizing variant or mutated Zace2 polypeptides. Studies that examine the interaction of Zace2 and competitive ACE inhibitors also can be employed to assay and characterize Zace2 polypeptides. Such assays are well known in the art. For a general reference, see Corvol *et al.*, *Meth. Enzymol.* 246:283 (1995). See also Williams *et al.*, *J. Biol. Chem.* 269:29430 (1994),

Sturrock *et al.*, *Biochem.* 35:9560 (1996), and Michaud *et al.*, *Molec. Pharmacol.* 51:1070 (1997).

As an illustration, a Zace2 variant can be tested for ACE activity using hippuryl-L-histidyl-L-leucine (Hip-His-Leu) as a substrate (see, for example, Sen *et al.*,  
5 *J. Biol. Chem.* 268:25748 (1993)). In one version of this assay, a solubilized test polypeptide is incubated in 0.4 M sodium borate buffer (pH 8.3) containing 300 mM sodium chloride for about 15 to 30 minutes at 37°C in the presence of varying concentrations of Hip-His-Leu (*e.g.*, 0.4 to 5 mM). The amount of His-Leu liberated by the test polypeptide is measured fluorometrically. Hip-His-Leu can also be used to  
10 identify Zace2 inhibitors by measuring the suppression of the cleavage of the substrate.

Other ACE substrates are known to those of skill in the art. For example, Isaac *et al.*, *Biochem. J.* 328:587 (1997), have shown that polypeptides having Lys/Arg-Arg at the C-terminus are high-affinity substrates for human tACE. Another useful substrate to measure ACE activity is [<sup>3</sup>H]benzyl-Phe-Ala-Pro (Howell *et al.*, *Am.*  
15 *J. Physiol.* 258:L188 (1990)).

Solid phase systems can also be used to identify a substrate or inhibitor of a Zace2 receptor polypeptide. For example, a Zace2 polypeptide or Zace2 fusion protein can be immobilized onto the surface of a receptor chip of a commercially available biosensor instrument (BIAcore, Biacore AB; Uppsala, Sweden). The use of  
20 this instrument is disclosed, for example, by Karlsson, *Immunol. Methods* 145:229 (1991), and Cunningham and Wells, *J. Mol. Biol.* 234:554 (1993).

In brief, a Zace2 polypeptide or fusion protein is covalently attached, using amine or sulfhydryl chemistry, to dextran fibers that are attached to gold film within a flow cell. A test sample is then passed through the cell. If a Zace2 substrate or  
25 inhibitor is present in the sample, it will bind to the immobilized polypeptide or fusion protein, causing a change in the refractive index of the medium, which is detected as a change in surface plasmon resonance of the gold film. This system allows the determination on- and off-rates, from which binding affinity can be calculated, and assessment of the stoichiometry of binding, as well as the kinetic effects of Zace2  
30 mutation. This system can also be used to examine antibody-antigen interactions, and the interactions of other complement/anti-complement pairs.

Accordingly, polypeptides of the present invention are useful as targets for identifying modulators of zinc protease activity. More particularly, Zace2 polypeptides are useful for screening or identifying new ACE inhibitors. The Zace2  
35 polypeptides can also be used as a basis for rational drug design of inhibitory molecules. These newly identified inhibitory molecules may be more specific or more potent than known ACE inhibitors. Moreover, Zace2 inhibitors may exhibit a more

favorable side effect profile than known ACE inhibitors. For example, Zace2 may contribute to certain unwanted side effects of ACE inhibitors, and as such, Zace2 would be useful to identify more specific ACE inhibitors.

In addition, inhibitory molecules identified using Zace2 polypeptides as  
5 a target may modulate different biological or physiological activities than known ACE inhibitors (e.g., the inhibitors may be effective for disorders other than those related to blood pressure and water and salt homeostasis). Zace2 inhibitors may provide broader inhibition than just ACE inhibition (for instance, these inhibitors may modulate many metalloprotease family members). Because Zace2 is more closely homologous to  
10 tACE than somatic ACE, Zace2 may permit selection of domain-specific inhibitors (those that inhibit the active site corresponding to the C domain of somatic ACE). Thus, a Zace2 inhibitor may specifically target angiotensin I and bradykinin-mediated effects, but have minimal or no effect on regulating hematopoiesis. Zace2 inhibitors may beneficially improve the status of patients with cardiovascular disease, and  
15 atherosclerotic vascular disease in particular, or renal disease, and diabetic nephropathy in particular. The effects of Zace2 inhibitors can be measured *in vitro* using cultured cells or *in vivo* by administering molecules of the claimed invention to the appropriate animal model.

The measurement of Zace2 enzyme activity can also be used for  
20 diagnosis. For example, the measurement of serum ACE activity levels provides useful information for the diagnosis of sarcoidosis and response to treatment (Studdy, *Lancet* 2(8104-5):1331 (1978)).

## 9. Production of Antibodies to Zace2 Proteins

25 Antibodies to Zace2 can be obtained, for example, using the product of a Zace2 expression vector or Zace2 isolated from a natural source as an antigen. Particularly useful anti-Zace2 antibodies "bind specifically" with Zace2. Antibodies are considered to be specifically binding if the antibodies exhibit at least one of the following two properties: (1) antibodies bind to Zace2 with a threshold level of binding  
30 activity, and (2) antibodies do not significantly cross-react with polypeptides related to Zace2.

With regard to the first characteristic, antibodies specifically bind if they bind to a Zace2 polypeptide, peptide or epitope with a binding affinity ( $K_d$ ) of  $10^6 M^{-1}$  or greater, preferably  $10^7 M^{-1}$  or greater, more preferably  $10^8 M^{-1}$  or greater, and most  
35 preferably  $10^9 M^{-1}$  or greater. The binding affinity of an antibody can be readily determined by one of ordinary skill in the art, for example, by Scatchard analysis

(Scatchard, *Ann. NY Acad. Sci.* 51:660 (1949)). With regard to the second characteristic, antibodies do not significantly cross-react with related polypeptide molecules, for example, if they detect Zace2, but not presently known polypeptides using a standard Western blot analysis. Examples of known related polypeptides are known angiotensin converting enzymes, such as human somatic ACE and tACE. Highly specifically binding anti-human Zace2 antibodies bind with human Zace2, but not murine Zace2, while highly specific anti-murine Zace2 antibodies bind with murine Zace2, but not with human Zace2.

Anti-Zace2 antibodies can be produced using antigenic Zace2 epitope-bearing peptides and polypeptides. Antigenic epitope-bearing peptides and polypeptides of the present invention contain a sequence of at least nine, or between 15 to about 30 amino acids contained within SEQ ID NOs:2, 6, or 9. However, peptides or polypeptides comprising a larger portion of an amino acid sequence of the invention, containing from 30 to 50 amino acids, or any length up to and including the entire amino acid sequence of a polypeptide of the invention, also are useful for inducing antibodies that bind with Zace2. It is desirable that the amino acid sequence of the epitope-bearing peptide is selected to provide substantial solubility in aqueous solvents (*i.e.*, the sequence includes relatively hydrophilic residues, while hydrophobic residues are preferably avoided). Moreover, amino acid sequences containing proline residues may be also be desirable for antibody production.

As an illustration, potential antigenic sites in human Zace2 were identified using the Jameson-Wolf method, Jameson and Wolf, *CABIOS* 4:181, (1988), as implemented by the PROTEAN program (version 3.14) of LASERGENE (DNASTAR; Madison, WI). Default parameters were used in this analysis.

The Jameson-Wolf method predicts potential antigenic determinants by combining six major subroutines for protein structural prediction. Briefly, the Hopp-Woods method, Hopp *et al.*, *Proc. Nat'l Acad. Sci. USA* 78:3824 (1981), was first used to identify amino acid sequences representing areas of greatest local hydrophilicity (parameter: seven residues averaged). In the second step, Emini's method, Emini *et al.*, *J. Virology* 55:836 (1985), was used to calculate surface probabilities (parameter: surface decision threshold  $(0.6) = 1$ ). Third, the Karplus-Schultz method, Karplus and Schultz, *Naturwissenschaften* 72:212 (1985), was used to predict backbone chain flexibility (parameter: flexibility threshold  $(0.2) = 1$ ). In the fourth and fifth steps of the analysis, secondary structure predictions were applied to the data using the methods of Chou-Fasman, Chou, "Prediction of Protein Structural Classes from Amino Acid Composition," in *Prediction of Protein Structure and the Principles of Protein Conformation*, Fasman (ed.), pages 549-586 (Plenum Press 1990), and Garnier-Robson,

Garnier *et al.*, *J. Mol. Biol.* 120:97 (1978) (Chou-Fasman parameters: conformation table = 64 proteins;  $\alpha$  region threshold = 103;  $\beta$  region threshold = 105; Garnier-Robson parameters:  $\alpha$  and  $\beta$  decision constants = 0). In the sixth subroutine, flexibility parameters and hydropathy/solvent accessibility factors were combined to determine a surface contour value, designated as the "antigenic index." Finally, a peak broadening function was applied to the antigenic index, which broadens major surface peaks by adding 20, 40, 60, or 80% of the respective peak value to account for additional free energy derived from the mobility of surface regions relative to interior regions. This calculation was not applied, however, to any major peak that resides in a helical region, since helical regions tend to be less flexible.

The results of this analysis indicated that the following amino acid sequences of SEQ ID NO:2 would provide suitable antigenic peptides: amino acids 19 to 26 ("antigenic peptide 1"), amino acids 56 to 69 ("antigenic peptide 2"), amino acids 135 to 141 ("antigenic peptide 3"), amino acids 169 to 181 ("antigenic peptide 4"), amino acids 196 to 210 ("antigenic peptide 5"), amino acids 196 to 236 ("antigenic peptide 6"), amino acids 285 to 295 ("antigenic peptide 7"), amino acids 426 to 436 ("antigenic peptide 8"), amino acids 491 to 502 ("antigenic peptide 9"), amino acids 534 to 552 ("antigenic peptide 10"), and amino acids 698 to 714 ("antigenic peptide 11"). The present invention contemplates the use of any one of antigenic peptides 1 to 11 to generate antibodies to Zace2. The present invention also contemplates polypeptides comprising at least one of antigenic peptides 1 to 11.

Similarly, Jameson-Wolf analysis of a murine Zace2 polypeptide indicated that the following amino acid sequences of SEQ ID NO:6 would provide suitable antigenic peptides: amino acids 19 to 26 ("antigenic peptide 12"), amino acids 33 to 39 ("antigenic peptide 13"), 54 to 64 ("antigenic peptide 14"), amino acids 74 to 81 ("antigenic peptide 15"), amino acids 134 to 140 ("antigenic peptide 16"), amino acids 156 to 161 ("antigenic peptide 17"), amino acids 204 to 215 ("antigenic peptide 18"), amino acids 427 to 435 ("antigenic peptide 19"), amino acids 491 to 499 ("antigenic peptide 20"), and amino acids 596 to 602 ("antigenic peptide 21"). The present invention contemplates the use of any one of antigenic peptides 12 to 21 to generate antibodies to murine Zace2. The present invention also contemplates polypeptides comprising at least one of antigenic peptides 12 to 21.

Polyclonal antibodies to recombinant Zace2 protein or to Zace2 isolated from natural sources can be prepared using methods well-known to those of skill in the art. See, for example, Green *et al.*, "Production of Polyclonal Antisera," in *Immunochemical Protocols* (Manson, ed.), pages 1-5 (Humana Press 1992), and Williams *et al.*, "Expression of foreign proteins in *E. coli* using plasmid vectors and



purification of specific polyclonal antibodies," in *DNA Cloning 2: Expression Systems, 2nd Edition*, Glover *et al.* (eds.), page 15 (Oxford University Press 1995). The immunogenicity of a Zace2 polypeptide can be increased through the use of an adjuvant, such as alum (aluminum hydroxide) or Freund's complete or incomplete adjuvant. Polypeptides useful for immunization also include fusion polypeptides, such as fusions of Zace2 or a portion thereof with an immunoglobulin polypeptide or with maltose binding protein. The polypeptide immunogen may be a full-length molecule or a portion thereof. If the polypeptide portion is "haptten-like," such portion may be advantageously joined or linked to a macromolecular carrier (such as keyhole limpet hemocyanin (KLH), bovine serum albumin (BSA) or tetanus toxoid) for immunization.

Although polyclonal antibodies are typically raised in animals such as horses, cows, dogs, chicken, rats, mice, rabbits, guinea pigs, goats, or sheep, an anti-Zace2 antibody of the present invention may also be derived from a subhuman primate antibody. General techniques for raising diagnostically and therapeutically useful antibodies in baboons may be found, for example, in Goldenberg *et al.*, international patent publication No. WO 91/11465, and in Losman *et al.*, *Int. J. Cancer* 46:310 (1990).

Alternatively, monoclonal anti-Zace2 antibodies can be generated. Rodent monoclonal antibodies to specific antigens may be obtained by methods known to those skilled in the art (see, for example, Kohler *et al.*, *Nature* 256:495 (1975), Coligan *et al.* (eds.), *Current Protocols in Immunology*, Vol. 1, pages 2.5.1-2.6.7 (John Wiley & Sons 1991) ["Coligan"], Picksley *et al.*, "Production of monoclonal antibodies against proteins expressed in *E. coli*," in *DNA Cloning 2: Expression Systems, 2nd Edition*, Glover *et al.* (eds.), page 93 (Oxford University Press 1995)).

Briefly, monoclonal antibodies can be obtained by injecting mice with a composition comprising a Zace2 gene product, verifying the presence of antibody production by removing a serum sample, removing the spleen to obtain B-lymphocytes, fusing the B-lymphocytes with myeloma cells to produce hybridomas, cloning the hybridomas, selecting positive clones, which produce antibodies to the antigen, culturing the clones that produce antibodies to the antigen, and isolating the antibodies from the hybridoma cultures.

In addition, an anti-Zace2 antibody of the present invention may be derived from a human monoclonal antibody. Human monoclonal antibodies are obtained from transgenic mice that have been engineered to produce specific human antibodies in response to antigenic challenge. In this technique, elements of the human heavy and light chain locus are introduced into strains of mice derived from embryonic stem cell lines that contain targeted disruptions of the endogenous heavy chain and light chain loci. The

transgenic mice can synthesize human antibodies specific for human antigens, and the mice can be used to produce human antibody-secreting hybridomas. Methods for obtaining human antibodies from transgenic mice are described, for example, by Green *et al.*, *Nature Genet.* 7:13 (1994), Lonberg *et al.*, *Nature* 368:856 (1994), and Taylor *et al.*,  
5 *Int. Immun.* 6:579 (1994).

Monoclonal antibodies can be isolated and purified from hybridoma cultures by a variety of well-established techniques. Such isolation techniques include affinity chromatography with Protein-A Sepharose, size-exclusion chromatography, and ion-exchange chromatography (see, for example, Coligan at pages 2.7.1-2.7.12 and  
10 pages 2.9.1-2.9.3; Baines *et al.*, "Purification of Immunoglobulin G (IgG)," in *Methods in Molecular Biology*, Vol. 10, pages 79-104 (The Humana Press, Inc. 1992)).

For particular uses, it may be desirable to prepare fragments of anti-Zace2 antibodies. Such antibody fragments can be obtained, for example, by proteolytic hydrolysis of the antibody. Antibody fragments can be obtained by pepsin  
15 or papain digestion of whole antibodies by conventional methods. As an illustration, antibody fragments can be produced by enzymatic cleavage of antibodies with pepsin to provide a 5S fragment denoted F(ab')<sub>2</sub>. This fragment can be further cleaved using a thiol reducing agent to produce 3.5S Fab' monovalent fragments. Optionally, the cleavage reaction can be performed using a blocking group for the sulfhydryl groups  
20 that result from cleavage of disulfide linkages. As an alternative, an enzymatic cleavage using pepsin produces two monovalent Fab fragments and an Fc fragment directly. These methods are described, for example, by Goldenberg, U.S. patent No. 4,331,647, Nisonoff *et al.*, *Arch Biochem. Biophys.* 89:230 (1960), Porter, *Biochem. J.* 73:119 (1959), Edelman *et al.*, in *Methods in Enzymology Vol. 1*, page 422 (Academic  
25 Press 1967), and by Coligan at pages 2.8.1-2.8.10 and 2.10.-2.10.4.

Other methods of cleaving antibodies, such as separation of heavy chains to form monovalent light-heavy chain fragments, further cleavage of fragments, or other enzymatic, chemical or genetic techniques may also be used, so long as the fragments bind to the antigen that is recognized by the intact antibody.

For example, Fv fragments comprise an association of V<sub>H</sub> and V<sub>L</sub> chains. This association can be noncovalent, as described by Inbar *et al.*, *Proc. Nat'l Acad. Sci. USA* 69:2659 (1972). Alternatively, the variable chains can be linked by an intermolecular disulfide bond or cross-linked by chemicals such as glutaraldehyde (see,  
30 for example, Sandhu, *Crit. Rev. Biotech.* 12:437 (1992)).

The Fv fragments may comprise V<sub>H</sub> and V<sub>L</sub> chains, which are connected by a peptide linker. These single-chain antigen binding proteins (scFv) are prepared by constructing a structural gene comprising DNA sequences encoding the V<sub>H</sub> and V<sub>L</sub>  
35

domains, which are connected by an oligonucleotide. The structural gene is inserted into an expression vector, which is subsequently introduced into a host cell, such as *E. coli*. The recombinant host cells synthesize a single polypeptide chain with a linker peptide bridging the two V domains. Methods for producing scFvs are described, for example, by Whitlow *et al.*, *Methods: A Companion to Methods in Enzymology* 2:97 (1991) (also see, Bird *et al.*, *Science* 242:423 (1988), Ladner *et al.*, U.S. Patent No. 4,946,778, Pack *et al.*, *Bio/Technology* 11:1271 (1993), and Sandhu, *supra*).

As an illustration, a scFV can be obtained by exposing lymphocytes to Zace2 polypeptide *in vitro*, and selecting antibody display libraries in phage or similar vectors (for instance, through use of immobilized or labeled Zace2 protein or peptide). Genes encoding polypeptides having potential Zace2 polypeptide binding domains can be obtained by screening random peptide libraries displayed on phage (phage display) or on bacteria, such as *E. coli*. Nucleotide sequences encoding the polypeptides can be obtained in a number of ways, such as through random mutagenesis and random polynucleotide synthesis. These random peptide display libraries can be used to screen for peptides, which interact with a known target, which can be a protein or polypeptide, such as a ligand or receptor, a biological or synthetic macromolecule, or organic or inorganic substances. Techniques for creating and screening such random peptide display libraries are known in the art (Ladner *et al.*, U.S. Patent No. 5,223,409, Ladner *et al.*, U.S. Patent No. 4,946,778, Ladner *et al.*, U.S. Patent No. 5,403,484, Ladner *et al.*, U.S. Patent No. 5,571,698, and Kay *et al.*, *Phage Display of Peptides and Proteins* (Academic Press, Inc. 1996)) and random peptide display libraries and kits for screening such libraries are available commercially, for instance from CLONTECH Laboratories, Inc. (Palo Alto, CA), Invitrogen Inc. (San Diego, CA), New England Biolabs, Inc. (Beverly, MA), and Pharmacia LKB Biotechnology Inc. (Piscataway, NJ). Random peptide display libraries can be screened using the Zace2 sequences disclosed herein to identify proteins, which bind to Zace2.

Another form of an antibody fragment is a peptide coding for a single complementarity-determining region (CDR). CDR peptides ("minimal recognition units") can be obtained by constructing genes encoding the CDR of an antibody of interest. Such genes are prepared, for example, by using the polymerase chain reaction to synthesize the variable region from RNA of antibody-producing cells (see, for example, Larrick *et al.*, *Methods: A Companion to Methods in Enzymology* 2:106 (1991), Courtenay-Luck, "Genetic Manipulation of Monoclonal Antibodies," in *Monoclonal Antibodies: Production, Engineering and Clinical Application*, Ritter *et al.* (eds.), page 166 (Cambridge University Press 1995), and Ward *et al.*, "Genetic

Manipulation and Expression of Antibodies," in *Monoclonal Antibodies: Principles and Applications*, Birch *et al.*, (eds.), page 137 (Wiley-Liss, Inc. 1995)).

Alternatively, an anti-Zace2 antibody may be derived from a "humanized" monoclonal antibody. Humanized monoclonal antibodies are produced by transferring mouse complementary determining regions from heavy and light variable chains of the mouse immunoglobulin into a human variable domain. Typical residues of human antibodies are then substituted in the framework regions of the murine counterparts. The use of antibody components derived from humanized monoclonal antibodies obviates potential problems associated with the immunogenicity of murine constant regions. General techniques for cloning murine immunoglobulin variable domains are described, for example, by Orlandi *et al.*, *Proc. Nat'l Acad. Sci. USA* 86:3833 (1989). Techniques for producing humanized monoclonal antibodies are described, for example, by Jones *et al.*, *Nature* 321:522 (1986), Carter *et al.*, *Proc. Nat'l Acad. Sci. USA* 89:4285 (1992), Sandhu, *Crit. Rev. Biotech.* 12:437 (1992), Singer *et al.*, *J. Immun.* 150:2844 (1993), Sudhir (ed.), *Antibody Engineering Protocols* (Humana Press, Inc. 1995), Kelley, "Engineering Therapeutic Antibodies," in *Protein Engineering: Principles and Practice*, Cleland *et al.* (eds.), pages 399-434 (John Wiley & Sons, Inc. 1996), and by Queen *et al.*, U.S. Patent No. 5,693,762 (1997).

Polyclonal anti-idiotypic antibodies can be prepared by immunizing animals with anti-Zace2 antibodies or antibody fragments, using standard techniques. See, for example, Green *et al.*, "Production of Polyclonal Antisera," in *Methods In Molecular Biology: Immunochemical Protocols*, Manson (ed.), pages 1-12 (Humana Press 1992). Also, see Coligan at pages 2.4.1-2.4.7. Alternatively, monoclonal anti-idiotypic antibodies can be prepared using anti-Zace2 antibodies or antibody fragments as immunogens with the techniques, described above. As another alternative, humanized anti-idiotypic antibodies or subhuman primate anti-idiotypic antibodies can be prepared using the above-described techniques. Methods for producing anti-idiotypic antibodies are described, for example, by Irie, U.S. Patent No. 5,208,146, Greene, *et al.*, U.S. Patent No. 5,637,677, and Varthakavi and Minocha, *J. Gen. Virol.* 77:1875 (1996).

#### **10. Use of Zace2 Nucleotide Sequences to Detect Gene Expression and Gene Structure**

Nucleic acid molecules can be used to detect the expression of a Zace2 gene in a biological sample. Suitable probe molecules include double-stranded nucleic acid molecules comprising the nucleotide sequence of SEQ ID NOs:1, 5, or 8, or a

portion thereof, as well as single-stranded nucleic acid molecules having the complement of the nucleotide sequence of SEQ ID NO:1, or a portion thereof. Probe molecules may be DNA, RNA, oligonucleotides, and the like. As used herein, the term "portion" refers to at least eight nucleotides to at least 20 or more nucleotides. Certain probes bind with regions of the *Zace2* gene that have a low sequence similarity to comparable regions in other proteins, such as other angiotensin converting enzymes.

In a basic assay, a single-stranded probe molecule is incubated with RNA, isolated from a biological sample, under conditions of temperature and ionic strength that promote base pairing between the probe and target *Zace2* RNA species. After separating unbound probe from hybridized molecules, the amount of hybrids is detected.

Well-established hybridization methods of RNA detection include northern analysis and dot/slot blot hybridization (see, for example, Ausubel (1995) at pages 4-1 to 4-27, and Wu *et al.* (eds.), "Analysis of Gene Expression at the RNA Level," in *Methods in Gene Biotechnology*, pages 225-239 (CRC Press, Inc. 1997)). Nucleic acid probes can be detectably labeled with radioisotopes such as  $^{32}\text{P}$  or  $^{35}\text{S}$ . Alternatively, *Zace2* RNA can be detected with a nonradioactive hybridization method (see, for example, Isaac (ed.), *Protocols for Nucleic Acid Analysis by Nonradioactive Probes* (Humana Press, Inc. 1993)). Typically, nonradioactive detection is achieved by enzymatic conversion of chromogenic or chemiluminescent substrates. Illustrative nonradioactive moieties include biotin, fluorescein, and digoxigenin.

*Zace2* oligonucleotide probes are also useful for *in vivo* diagnosis. As an illustration,  $^{18}\text{F}$ -labeled oligonucleotides can be administered to a subject and visualized by positron emission tomography (Tavitian *et al.*, *Nature Medicine* 4:467 (1998)).

Numerous diagnostic procedures take advantage of the polymerase chain reaction (PCR) to increase sensitivity of detection methods. Standard techniques for performing PCR are well-known (see, generally, Mathew (ed.), *Protocols in Human Molecular Genetics* (Humana Press, Inc. 1991), White (ed.), *PCR Protocols: Current Methods and Applications* (Humana Press, Inc. 1993), Cotter (ed.), *Molecular Diagnosis of Cancer* (Humana Press, Inc. 1996), Hanausek and Walaszek (eds.), *Tumor Marker Protocols* (Humana Press, Inc. 1998), Lo (ed.), *Clinical Applications of PCR* (Humana Press, Inc. 1998), and Meltzer (ed.), *PCR in Bioanalysis* (Humana Press, Inc. 1998)).

PCR primers can be designed to amplify a portion of the *Zace2* gene that has a low sequence similarity to a comparable region in other proteins, such as other angiotensin converting enzymes.

One variation of PCR for diagnostic assays is reverse transcriptase-PCR (RT-PCR). In the RT-PCR technique, RNA is isolated from a biological sample, reverse transcribed to cDNA, and the cDNA is incubated with *Zace2* primers (see, for example, Wu *et al.* (eds.), "Rapid Isolation of Specific cDNAs or Genes by PCR," in *Methods in Gene Biotechnology*, pages 15-28 (CRC Press, Inc. 1997)). PCR is then performed and the products are analyzed using standard techniques.

As an illustration, RNA is isolated from biological sample using, for example, the guanidium-thiocyanate cell lysis procedure described above. Alternatively, a solid-phase technique can be used to isolate mRNA from a cell lysate. A reverse transcription reaction can be primed with the isolated RNA using random oligonucleotides, short homopolymers of dT, or *Zace2* anti-sense oligomers. Oligo-dT primers offer the advantage that various mRNA nucleotide sequences are amplified that can provide control target sequences. *Zace2* sequences are amplified by the polymerase chain reaction using two flanking oligonucleotide primers that are typically 20 bases in length.

PCR amplification products can be detected using a variety of approaches. For example, PCR products can be fractionated by gel electrophoresis, and visualized by ethidium bromide staining. Alternatively, fractionated PCR products can be transferred to a membrane, hybridized with a detectably-labeled *Zace2* probe, and examined by autoradiography. Additional alternative approaches include the use of digoxigenin-labeled deoxyribonucleic acid triphosphates to provide chemiluminescence detection, and the C-TRAK colorimetric assay.

Another approach for detection of *Zace2* expression is cycling probe technology (CPT), in which a single-stranded DNA target binds with an excess of DNA-RNA-DNA chimeric probe to form a complex, the RNA portion is cleaved with RNAase H, and the presence of cleaved chimeric probe is detected (see, for example, Beggs *et al.*, *J. Clin. Microbiol.* 34:2985 (1996), Bekkaoui *et al.*, *Biotechniques* 20:240 (1996)). Alternative methods for detection of *Zace2* sequences can utilize approaches such as nucleic acid sequence-based amplification (NASBA), cooperative amplification of templates by cross-hybridization (CATCH), and the ligase chain reaction (LCR) (see, for example, Marshall *et al.*, U.S. Patent No. 5,686,272 (1997), Dyer *et al.*, *J. Virol. Methods* 60:161 (1996), Ehricht *et al.*, *Eur. J. Biochem.* 243:358 (1997), and Chadwick *et al.*, *J. Virol. Methods* 70:59 (1998)). Other standard methods are known to those of skill in the art.

*Zace2* probes and primers can also be used to detect and to localize *Zace2* gene expression in tissue samples. Methods for such *in situ* hybridization are well-known to those of skill in the art (see, for example, Choo (ed.), *In Situ*

*Hybridization Protocols* (Humana Press, Inc. 1994), Wu *et al.* (eds.), "Analysis of Cellular DNA or Abundance of mRNA by Radioactive *In Situ* Hybridization (RISH)," in *Methods in Gene Biotechnology*, pages 259-278 (CRC Press, Inc. 1997), and Wu *et al.* (eds.), "Localization of DNA or Abundance of mRNA by Fluorescence *In Situ* Hybridization (RISH)," in *Methods in Gene Biotechnology*, pages 279-289 (CRC Press, Inc. 1997)). Various additional diagnostic approaches are well-known to those of skill in the art (see, for example, Mathew (ed.), *Protocols in Human Molecular Genetics* (Humana Press, Inc. 1991), Coleman and Tsongalis, *Molecular Diagnostics* (Humana Press, Inc. 1996), and Elles, *Molecular Diagnosis of Genetic Diseases* (Humana Press, Inc., 1996)). Suitable test samples include blood, urine, saliva, tissue biopsy, and autopsy material.

Clinically significant polymorphisms of the human *ACE* gene have been discovered (see, for example, Matsusaka and Ichikawa, *Annu. Rev. Physiol.* 59:395 (1997)). A polymorphism associated with intron 16 is associated with plasma and intracellular levels of ACE, as well as increased risk of myocardial infarction. *ACE* polymorphisms are also associated with progression to chronic renal failure in IgA nephropathy, and diabetic nephropathy (Marre *et al.*, *Diabetes* 43:384 (1994); Yoshida *et al.*, *J. Clin. Invest.* 96:2162 (1995)). Other *ACE* gene mutations are associated with the risk of developing cardiovascular disease (Raynolds and Perryman, U.S. Patent No. 5,800, 990).

Nucleic acid molecules comprising *Zace2* nucleotide sequences can also be used to determine whether a subject's chromosomes contain a mutation in the *Zace2* gene. Detectable chromosomal aberrations at the *Zace2* gene locus include, but are not limited to, aneuploidy, gene copy number changes, insertions, deletions, restriction site changes and rearrangements. Of particular interest are genetic alterations that inactivate the *Zace2* gene.

Aberrations associated with the *Zace2* locus can be detected using nucleic acid molecules of the present invention by employing molecular genetic techniques, such as restriction fragment length polymorphism analysis, short tandem repeat analysis employing PCR techniques, amplification-refractory mutation system analysis, single-strand conformation polymorphism detection, RNase cleavage methods, denaturing gradient gel electrophoresis, fluorescence-assisted mismatch analysis, and other genetic analysis techniques known in the art (see, for example, Mathew (ed.), *Protocols in Human Molecular Genetics* (Humana Press, Inc. 1991), Marian, *Chest* 108:255 (1995), Coleman and Tsongalis, *Molecular Diagnostics* (Humana Press, Inc. 1996), Elles (ed.) *Molecular Diagnosis of Genetic Diseases* (Humana Press, Inc. 1996), Landegren (ed.), *Laboratory Protocols for Mutation*

*Detection* (Oxford University Press 1996), Birren *et al.* (eds.), *Genome Analysis, Vol. 2: Detecting Genes* (Cold Spring Harbor Laboratory Press 1998), Dracopoli *et al.* (eds.), *Current Protocols in Human Genetics* (John Wiley & Sons 1998), and Richards and Ward, "Molecular Diagnostic Testing," in *Principles of Molecular Medicine*, pages 83-88 (Humana Press, Inc. 1998)).

The protein truncation test is also useful for detecting the inactivation of a gene in which translation-terminating mutations produce only portions of the encoded protein (see, for example, Stoppa-Lyonnet *et al.*, *Blood* 91:3920 (1998)). According to this approach, RNA is isolated from a biological sample, and used to synthesize cDNA. PCR is then used to amplify the *Zace2* target sequence and to introduce an RNA polymerase promoter, a translation initiation sequence, and an in-frame ATG triplet. PCR products are transcribed using an RNA polymerase, and the transcripts are translated *in vitro* with a T7-coupled reticulocyte lysate system. The translation products are then fractionated by SDS-PAGE to determine the lengths of the translation products. The protein truncation test is described, for example, by Dracopoli *et al.* (eds.), *Current Protocols in Human Genetics*, pages 9.11.1 - 9.11.18 (John Wiley & Sons 1998).

The *Zace2* gene resides on the human X chromosome. Further localization studies revealed that the *Zace2* gene resides at Xp22.1. This region is also associated with X-linked forms of mental retardation and infantile spasm syndrome.

The present invention contemplates kits for performing a diagnostic assay for *Zace2* gene expression or to detect mutations in the *Zace2* gene. Such kits comprise nucleic acid probes, such as double-stranded nucleic acid molecules comprising the nucleotide sequence of SEQ ID NO:1, or a portion thereof, as well as single-stranded nucleic acid molecules having the complement of the nucleotide sequence of SEQ ID NO:1, or a portion thereof. Probe molecules may be DNA, RNA, oligonucleotides, and the like. Kits may comprise nucleic acid primers for performing PCR.

Such a kit can contain all the necessary elements to perform a nucleic acid diagnostic assay described above. A kit will comprise at least one container comprising a *Zace2* probe or primer. The kit may also comprise a second container comprising one or more reagents capable of indicating the presence of *Zace2* sequences. Examples of such indicator reagents include detectable labels such as radioactive labels, fluorochromes, chemiluminescent agents, and the like. A kit may also comprise a means for conveying to the user that the *Zace2* probes and primers are used to detect *Zace2* gene expression. For example, written instructions may state that the enclosed nucleic acid molecules can be used to detect either a nucleic acid molecule that encodes *Zace2*, or a nucleic acid molecule having a nucleotide sequence that is



complementary to a *Zace2*-encoding nucleotide sequence. The written material can be applied directly to a container, or the written material can be provided in the form of a packaging insert.

5     **11. Use of Anti-Zace2 Antibodies to Detect Zace2**

The present invention contemplates the use of anti-Zace2 antibodies to screen biological samples *in vitro* for the presence of Zace2. In one type of *in vitro* assay, anti-Zace2 antibodies are used in liquid phase. For example, the presence of Zace2 in a biological sample can be tested by mixing the biological sample with a trace amount of  
10     labeled Zace2 and an anti-Zace2 antibody under conditions that promote binding between Zace2 and its antibody. Complexes of Zace2 and anti-Zace2 in the sample can be separated from the reaction mixture by contacting the complex with an immobilized protein, which binds with the antibody, such as an Fc antibody or *Staphylococcus* protein A. The concentration of Zace2 in the biological sample will be inversely proportional to  
15     the amount of labeled Zace2 bound to the antibody and directly related to the amount of free labeled Zace2. Illustrative biological samples include blood, urine, saliva, tissue biopsy, and autopsy material.

Alternatively, *in vitro* assays can be performed in which anti-Zace2 antibody is bound to a solid-phase carrier. For example, antibody can be attached to a  
20     polymer, such as aminodextran, in order to link the antibody to an insoluble support such as a polymer-coated bead, a plate or a tube. Other suitable *in vitro* assays will be readily apparent to those of skill in the art.

In another approach, anti-Zace2 antibodies can be used to detect Zace2 in tissue sections prepared from a biopsy specimen. Such immunochemical detection can be  
25     used to determine the relative abundance of Zace2 and to determine the distribution of Zace2 in the examined tissue. General immunochemistry techniques are well established (see, for example, Ponder, "Cell Marking Techniques and Their Application," in *Mammalian Development: A Practical Approach*, Monk (ed.), pages 115-38 (IRL Press 1987), Coligan at pages 5.8.1-5.8.8, Ausubel (1995) at pages 14.6.1 to 14.6.13 (Wiley  
30     Interscience 1990), and Manson (ed.), *Methods In Molecular Biology, Vol. 10: Immunochemical Protocols* (The Humana Press, Inc. 1992)).

Immunochemical detection can be performed by contacting a biological sample with an anti-Zace2 antibody, and then contacting the biological sample with a detectably labeled molecule, which binds to the antibody. For example, the detectably  
35     labeled molecule can comprise an antibody moiety that binds to anti-Zace2 antibody. Alternatively, the anti-Zace2 antibody can be conjugated with avidin/streptavidin (or

biotin) and the detectably labeled molecule can comprise biotin (or avidin/streptavidin). Numerous variations of this basic technique are well-known to those of skill in the art.

Alternatively, an anti-Zace2 antibody can be conjugated with a detectable label to form an anti-Zace2 immunoconjugate. Suitable detectable labels include, for example, a radioisotope, a fluorescent label, a chemiluminescent label, an enzyme label, a bioluminescent label or colloidal gold. Methods of making and detecting such detectably-labeled immunoconjugates are well-known to those of ordinary skill in the art, and are described in more detail below.

The detectable label can be a radioisotope that is detected by autoradiography. Isotopes that are particularly useful for the purpose of the present invention are  $^3\text{H}$ ,  $^{125}\text{I}$ ,  $^{131}\text{I}$ ,  $^{35}\text{S}$  and  $^{14}\text{C}$ .

Anti-Zace2 immunoconjugates can also be labeled with a fluorescent compound. The presence of a fluorescently-labeled antibody is determined by exposing the immunoconjugate to light of the proper wavelength and detecting the resultant fluorescence. Fluorescent labeling compounds include fluorescein isothiocyanate, rhodamine, phycoerytherin, phycocyanin, allophycocyanin, o-phthaldehyde and fluorescamine.

Alternatively, anti-Zace2 immunoconjugates can be detectably labeled by coupling an antibody component to a chemiluminescent compound. The presence of the chemiluminescent-tagged immunoconjugate is determined by detecting the presence of luminescence that arises during the course of a chemical reaction. Examples of chemiluminescent labeling compounds include luminol, isoluminol, an aromatic acridinium ester, an imidazole, an acridinium salt and an oxalate ester.

Similarly, a bioluminescent compound can be used to label anti-Zace2 immunoconjugates of the present invention. Bioluminescence is a type of chemiluminescence found in biological systems in which a catalytic protein increases the efficiency of the chemiluminescent reaction. The presence of a bioluminescent protein is determined by detecting the presence of luminescence. Bioluminescent compounds that are useful for labeling include luciferin, luciferase and aequorin.

Alternatively, anti-Zace2 immunoconjugates can be detectably labeled by linking an anti-Zace2 antibody component to an enzyme. When the anti-Zace2-enzyme conjugate is incubated in the presence of the appropriate substrate, the enzyme moiety reacts with the substrate to produce a chemical moiety, which can be detected, for example, by spectrophotometric, fluorometric or visual means. Examples of enzymes that can be used to detectably label polyspecific immunoconjugates include  $\beta$ -galactosidase, glucose oxidase, peroxidase and alkaline phosphatase.

Those of skill in the art will know of other suitable labels, which can be employed in accordance with the present invention. The binding of marker moieties to

anti-Zace2 antibodies can be accomplished using standard techniques known to the art. Typical methodology in this regard is described by Kennedy *et al.*, *Clin. Chim. Acta* 70:1 (1976), Schurs *et al.*, *Clin. Chim. Acta* 81:1 (1977), Shih *et al.*, *Int'l J. Cancer* 46:1101 (1990), Stein *et al.*, *Cancer Res.* 50:1330 (1990), and Coligan, *supra*.

5           Moreover, the convenience and versatility of immunochemical detection can be enhanced by using anti-Zace2 antibodies that have been conjugated with avidin, streptavidin, and biotin (see, for example, Wilchek *et al.* (eds.), "Avidin-Biotin Technology," *Methods In Enzymology*, Vol. 184 (Academic Press 1990), and Bayer *et al.*, "Immunochemical Applications of Avidin-Biotin Technology," in *Methods In Molecular*  
10 *Biology*, Vol. 10, Manson (ed.), pages 149-162 (The Humana Press, Inc. 1992).

          Methods for performing immunoassays are well-established. See, for example, Cook and Self, "Monoclonal Antibodies in Diagnostic Immunoassays," in *Monoclonal Antibodies: Production, Engineering, and Clinical Application*, Ritter and Ladyman (eds.), pages 180-208, (Cambridge University Press, 1995), Perry, "The Role of  
15 Monoclonal Antibodies in the Advancement of Immunoassay Technology," in *Monoclonal Antibodies: Principles and Applications*, Birch and Lennox (eds.), pages 107-120 (Wiley-Liss, Inc. 1995), and Diamandis, *Immunoassay* (Academic Press, Inc. 1996).

          The present invention also contemplates kits for performing an  
20 immunological diagnostic assay for Zace2 gene expression. Such kits comprise at least one container comprising an anti-Zace2 antibody, or antibody fragment. A kit may also comprise a second container comprising one or more reagents capable of indicating the presence of Zace2 antibody or antibody fragments. Examples of such indicator reagents include detectable labels such as a radioactive label, a fluorescent label, a  
25 chemiluminescent label, an enzyme label, a bioluminescent label, colloidal gold, and the like. A kit may also comprise a means for conveying to the user that Zace2 antibodies or antibody fragments are used to detect Zace2 protein. For example, written instructions may state that the enclosed antibody or antibody fragment can be used to detect Zace2. The written material can be applied directly to a container, or the written  
30 material can be provided in the form of a packaging insert.

## 12. Therapeutic Uses of Polypeptides Having Zace2 Activity

          The present invention includes the use of proteins, polypeptides, and peptides having Zace2 activity (such as Zace2 polypeptides (*e.g.*, soluble forms of  
35 Zace2), Zace2 analogs (*e.g.*, anti-Zace2 anti-idiotypic antibodies), and Zace2 fusion proteins) to a subject who lacks an adequate amount of this polypeptide. In contrast,

Zace2 antagonists (e.g., anti-Zace2 antibodies) can be used to treat a subject who produces an excess of Zace2. Either human or murine Zace2 protein can be used for such methods.

5 The kallikrein-kinin (contact) system modulates the renin-angiotensin-aldosterone system, prostaglandins, vasopressins, sodium-water balance, renal hemodynamics, and blood pressure. Stadnicki *et al.*, *FASEB J.* 12:325 (1998), have shown that a reversible inhibitor of plasma kallikrein decreased chronic intestinal inflammation in an experimental model relevant to Crohn's disease. One of the actions of kallikrein is to cleave high molecular weight kininogen to produce bradykinin, a  
10 peptide that enhances vasodilation, increases vascular permeability, and influences intestinal motility and electrolyte secretion (see, for example, Bhoola *et al.*, *Pharmacol. Rev.* 44:1 (1992)). The inhibition of kallikrein by the reversible inhibitor, therefore, should decrease bradykinin activity levels, which is consistent with evidence that kinins mediate gastrointestinal inflammation associated with inflammatory bowel disease,  
15 such as Crohn's disease (see, for example, Bachvarov *et al.*, *Gastroenterology* 115:1045 (1998)).

ACE also decreases bradykinin activity by cleaving the peptide. Accordingly, decreased ACE activity should be correlated with increased bradykinin activity. Studies have shown that serum ACE activity is significantly lowered in  
20 certain patients who have active Crohn's disease (see, for example, Silverstein *et al.*, *Am. J. Clin. Pathol.* 75:175 (1981); Sommer *et al.*, *Enzyme* 35:181 (1986)). Taken together, these observations indicate that ACE can be used to treat conditions associated with inflammation, such as inflammatory bowel disease.

The present invention therefore includes the use of polypeptides having  
25 Zace2 activity (e.g., Zace2 polypeptides, functional fragments of Zace2, anti-Zace2 anti-idiotype antibodies, *etc.*) to treat an inflammatory bowel disease (e.g., Crohn's disease and ulcerative colitis). More generally, the present invention includes the use of polypeptides having Zace2 activity to treat diseases associated with inflammation, such as arthritis and enterocolitis, two conditions that have been treated with a  
30 kallikrein inhibitor (see, for example, DeLa Cadena *et al.*, *FASEB J.* 9:446 (1995); Stadnicki *et al.*, *Dig. Dis. Sci.* 41:912 (1996)). Methods for identification of subjects suitable for such treatment are well known to those of skill in the art (see, for example, Rakel (ed.), *Conn's 1999 Current Therapy* (W.B. Saunders Company 1999)).

As described above, northern analysis indicates that human Zace2 is  
35 expressed by testicular tissue at high levels. This observation in combination with the single putative catalytic domain of Zace2 indicates that Zace2 shares properties with the known testicular form of ACE. Studies in murine models have shown that testicular

ACE is required for fertility (see, for example, Bernstein *et al.*, *Am. J. Cardiol.* 82 (10A):64 (1998), Hagaman *et al.*, *Proc. Nat'l Acad. Sci. USA* 95:2552 (1998), and Ramaraj *et al.*, *J. Clin. Invest.* 102:371 (1998)). Accordingly, the present invention includes the use of Zace2 agonists to treat infertility, and the use of Zace2 antagonists to induce infertility.

Generally, the dosage of administered Zace2 (or Zace2 analog or fusion protein) will vary depending upon such factors as the patient's age, weight, height, sex, general medical condition and previous medical history. Typically, it is desirable to provide the recipient with a dosage of Zace2, which is in the range of from about 1 pg/kg to 10 mg/kg (amount of agent/body weight of patient), although a lower or higher dosage also may be administered as circumstances dictate.

Administration of a molecule having Zace2 activity to a subject can be intravenous, intraarterial, intraperitoneal, intramuscular, subcutaneous, intrapleural, intrathecal, by perfusion through a regional catheter, or by direct intralesional injection. Regional administration is particularly useful for treatment of an inflammatory bowel disease. When administering therapeutic proteins by injection, the administration may be by continuous infusion or by single or multiple boluses.

Additional routes of administration include oral, mucosal-membrane, pulmonary, and transcutaneous. Oral delivery is suitable for polyester microspheres, zein microspheres, proteinoid microspheres, polycyanoacrylate microspheres, and lipid-based systems (see, for example, DiBase and Morrel, "Oral Delivery of Microencapsulated Proteins," in *Protein Delivery: Physical Systems*, Sanders and Hendren (eds.), pages 255-288 (Plenum Press 1997)). The feasibility of an intranasal delivery is exemplified by such a mode of insulin administration (see, for example, Hinchcliffe and Illum, *Adv. Drug Deliv. Rev.* 35:199 (1999)). Dry or liquid particles comprising Zace2 can be prepared and inhaled with the aid of dry-powder dispersers, liquid aerosol generators, or nebulizers (e.g., Pettit and Gombotz, *TIBTECH* 16:343 (1998); Patton *et al.*, *Adv. Drug Deliv. Rev.* 35:235 (1999)). This approach is illustrated by the AERX diabetes management system, which is a hand-held electronic inhaler that delivers aerosolized insulin into the lungs. Studies have shown that proteins as large as 48,000 kDa have been delivered across skin at therapeutic concentrations with the aid of low-frequency ultrasound, which illustrates the feasibility of transcutaneous administration (Mitragotri *et al.*, *Science* 269:850 (1995)). Transdermal delivery using electroporation provides another means to administer a molecule having Zace2 activity (Potts *et al.*, *Pharm. Biotechnol.* 10:213 (1997)).

A pharmaceutical composition comprising a protein, polypeptide, or peptide having Zace2 activity can be formulated according to known methods to

prepare pharmaceutically useful compositions, whereby the therapeutic proteins are combined in a mixture with a pharmaceutically acceptable carrier. A composition is said to be a "pharmaceutically acceptable carrier" if its administration can be tolerated by a recipient patient. Sterile phosphate-buffered saline is one example of a pharmaceutically acceptable carrier. Other suitable carriers are well-known to those in the art. See, for example, Gennaro (ed.), *Remington's Pharmaceutical Sciences*, 19th Edition (Mack Publishing Company 1995).

For purposes of therapy, molecules having Zace2 activity and a pharmaceutically acceptable carrier are administered to a patient in a therapeutically effective amount. A combination of a protein, polypeptide, or peptide having Zace2 activity and a pharmaceutically acceptable carrier is said to be administered in a "therapeutically effective amount" if the amount administered is physiologically significant. An agent is physiologically significant if its presence results in a detectable change in the physiology of a recipient patient. For example, common symptoms of Crohn's disease include chronic diarrhea with abdominal pain, fever, anorexia, weight loss, and a right lower quadrant mass. An agent used to treat Crohn's disease is physiologically significant if its presence alleviates at least one of these symptoms.

A pharmaceutical composition comprising Zace2 (or Zace2 analog or fusion protein) can be furnished in liquid form, in an aerosol, or in solid form. Liquid forms, are illustrated by injectable solutions and oral suspensions. Exemplary solid forms include capsules, tablets, and controlled-release forms. The latter form is illustrated by miniosmotic pumps and implants (Bremer *et al.*, *Pharm. Biotechnol.* 10:239 (1997); Ranade, "Implants in Drug Delivery," in *Drug Delivery Systems*, Ranade and Hollinger (eds.), pages 95-123 (CRC Press 1995); Bremer *et al.*, "Protein Delivery with Infusion Pumps," in *Protein Delivery: Physical Systems*, Sanders and Hendren (eds.), pages 239-254 (Plenum Press 1997); Yewey *et al.*, "Delivery of Proteins from a Controlled Release Injectable Implant," in *Protein Delivery: Physical Systems*, Sanders and Hendren (eds.), pages 93-117 (Plenum Press 1997)).

Liposomes provide one means to deliver therapeutic polypeptides to a subject intravenously, intraperitoneally, intrathecally, intramuscularly, subcutaneously, or via oral administration, inhalation, or intranasal administration. Liposomes are microscopic vesicles that consist of one or more lipid bilayers surrounding aqueous compartments (see, generally, Bakker-Woudenberg *et al.*, *Eur. J. Clin. Microbiol. Infect. Dis.* 12 (Suppl. 1):S61 (1993), Kim, *Drugs* 46:618 (1993), and Ranade, "Site-Specific Drug Delivery Using Liposomes as Carriers," in *Drug Delivery Systems*, Ranade and Hollinger (eds.), pages 3-24 (CRC Press 1995)). Liposomes are similar in composition to cellular membranes and as a result, liposomes can be administered

safely and are biodegradable. Depending on the method of preparation, liposomes may be unilamellar or multilamellar, and liposomes can vary in size with diameters ranging from 0.02  $\mu\text{m}$  to greater than 10  $\mu\text{m}$ . A variety of agents can be encapsulated in liposomes: hydrophobic agents partition in the bilayers and hydrophilic agents partition within the inner aqueous space(s) (see, for example, Machy *et al.*, *Liposomes In Cell Biology And Pharmacology* (John Libbey 1987), and Ostro *et al.*, *American J. Hosp. Pharm.* 46:1576 (1989)). Moreover, it is possible to control the therapeutic availability of the encapsulated agent by varying liposome size, the number of bilayers, lipid composition, as well as the charge and surface characteristics of the liposomes.

Liposomes can adsorb to virtually any type of cell and then slowly release the encapsulated agent. Alternatively, an absorbed liposome may be endocytosed by cells that are phagocytic. Endocytosis is followed by intralysosomal degradation of liposomal lipids and release of the encapsulated agents (Scherphof *et al.*, *Ann. N.Y. Acad. Sci.* 446:368 (1985)). After intravenous administration, small liposomes (0.1 to 1.0  $\mu\text{m}$ ) are typically taken up by cells of the reticuloendothelial system, located principally in the liver and spleen, whereas liposomes larger than 3.0  $\mu\text{m}$  are deposited in the lung. This preferential uptake of smaller liposomes by the cells of the reticuloendothelial system has been used to deliver chemotherapeutic agents to macrophages and to tumors of the liver.

The reticuloendothelial system can be circumvented by several methods including saturation with large doses of liposome particles, or selective macrophage inactivation by pharmacological means (Claassen *et al.*, *Biochim. Biophys. Acta* 802:428 (1984)). In addition, incorporation of glycolipid- or polyethylene glycol-derivatized phospholipids into liposome membranes has been shown to result in a significantly reduced uptake by the reticuloendothelial system (Allen *et al.*, *Biochim. Biophys. Acta* 1068:133 (1991); Allen *et al.*, *Biochim. Biophys. Acta* 1150:9 (1993)).

Liposomes can also be prepared to target particular cells or organs by varying phospholipid composition or by inserting receptors or ligands into the liposomes. For example, liposomes, prepared with a high content of a nonionic surfactant, have been used to target the liver (Hayakawa *et al.*, Japanese Patent 04-244,018; Kato *et al.*, *Biol. Pharm. Bull.* 16:960 (1993)). These formulations were prepared by mixing soybean phosphatidylcholine,  $\alpha$ -tocopherol, and ethoxylated hydrogenated castor oil (HCO-60) in methanol, concentrating the mixture under vacuum, and then reconstituting the mixture with water. A liposomal formulation of dipalmitoylphosphatidylcholine (DPPC) with a soybean-derived sterylglucoside mixture (SG) and cholesterol (Ch) has also been shown to target the liver (Shimizu *et al.*, *Biol. Pharm. Bull.* 20:881 (1997)).

Alternatively, various targeting ligands can be bound to the surface of the liposome, such as antibodies, antibody fragments, carbohydrates, vitamins, and transport protein. For example, liposomes can be modified with branched type galactosyllipid derivatives to target asialoglycoprotein (galactose) receptors, which are exclusively expressed on the surface of liver cells (Kato and Sugiyama, *Crit. Rev. Ther. Drug Carrier Syst.* 14:287 (1997); Murahashi *et al.*, *Biol. Pharm. Bull.* 20:259 (1997)). Similarly, Wu *et al.*, *Hepatology* 27:772 (1998), have shown that labeling liposomes with asialofetuin led to a shortened liposome plasma half-life and greatly enhanced uptake of asialofetuin-labeled liposome by hepatocytes. On the other hand, hepatic accumulation of liposomes comprising branched type galactosyllipid derivatives can be inhibited by preinjection of asialofetuin (Murahashi *et al.*, *Biol. Pharm. Bull.* 20:259 (1997)). Polyacornitylated human serum albumin liposomes provide another approach for targeting liposomes to liver cells (Kamps *et al.*, *Proc. Nat'l Acad. Sci. USA* 94:11681 (1997)). Moreover, Geho, *et al.* U.S. Patent No. 4,603,044, describe a hepatocyte-directed liposome vesicle delivery system, which has specificity for hepatobiliary receptors associated with the specialized metabolic cells of the liver.

In a more general approach to tissue targeting, target cells are prelabeled with biotinylated antibodies specific for a ligand expressed by the target cell (Harasym *et al.*, *Adv. Drug Deliv. Rev.* 32:99 (1998)). After plasma elimination of free antibody, streptavidin-conjugated liposomes are administered. In another approach, targeting antibodies are directly attached to liposomes (Harasym *et al.*, *Adv. Drug Deliv. Rev.* 32:99 (1998)).

Polypeptides having Zacc2 activity can be encapsulated within liposomes using standard techniques of protein microencapsulation (see, for example, Anderson *et al.*, *Infect. Immun.* 31:1099 (1981), Anderson *et al.*, *Cancer Res.* 50:1853 (1990), and Cohen *et al.*, *Biochim. Biophys. Acta* 1063:95 (1991), Alving *et al.* "Preparation and Use of Liposomes in Immunological Studies," in *Liposome Technology*, 2nd Edition, Vol. III, Gregoriadis (ed.), page 317 (CRC Press 1993), Wassef *et al.*, *Meth. Enzymol.* 149:124 (1987)). As noted above, therapeutically useful liposomes may contain a variety of components. For example, liposomes may comprise lipid derivatives of poly(ethylene glycol) (Allen *et al.*, *Biochim. Biophys. Acta* 1150:9 (1993)).

Degradable polymer microspheres have been designed to maintain high systemic levels of therapeutic proteins. Microspheres are prepared from degradable polymers such as poly(lactide-co-glycolide) (PLG), polyanhydrides, poly (ortho esters), nonbiodegradable ethylvinyl acetate polymers, in which proteins are entrapped in the polymer (Gombotz and Pettit, *Bioconjugate Chem.* 6:332 (1995); Ranade, "Role of



Polymers in Drug Delivery," in *Drug Delivery Systems*, Ranade and Hollinger (eds.), pages 51-93 (CRC Press 1995); Roskos and Maskiewicz, "Degradable Controlled Release Systems Useful for Protein Delivery," in *Protein Delivery: Physical Systems*, Sanders and Hendren (eds.), pages 45-92 (Plenum Press 1997); Bartus *et al.*, *Science* 281:1161 (1998); Putney and Burke, *Nature Biotechnology* 16:153 (1998); Putney, *Curr. Opin. Chem. Biol.* 2:548 (1998)). Polyethylene glycol (PEG)-coated nanospheres can also provide carriers for intravenous administration of therapeutic proteins (see, for example, Gref *et al.*, *Pharm. Biotechnol.* 10:167 (1997)).

The present invention also contemplates chemically modified polypeptides having Zace2 activity and Zace2 antagonists, in which a polypeptide is linked with a polymer, as discussed above.

Other dosage forms can be devised by those skilled in the art, as shown, for example, by Ansel and Popovich, *Pharmaceutical Dosage Forms and Drug Delivery Systems*, 5<sup>th</sup> Edition (Lea & Febiger 1990), Gennaro (ed.), *Remington's Pharmaceutical Sciences*, 19<sup>th</sup> Edition (Mack Publishing Company 1995), and by Ranade and Hollinger, *Drug Delivery Systems* (CRC Press 1996).

As an illustration, pharmaceutical compositions may be supplied as a kit comprising a container that comprises a molecule having Zace2 activity or a Zace2 antagonist (e.g., an antibody or antibody fragment that binds a Zace2 polypeptide). Therapeutic polypeptides can be provided in the form of an injectable solution for single or multiple doses, or as a sterile powder that will be reconstituted before injection. Alternatively, such a kit can include a dry-powder disperser, liquid aerosol generator, or nebulizer for administration of a therapeutic polypeptide. Such a kit may further comprise written information on indications and usage of the pharmaceutical composition. Moreover, such information may include a statement that the Zace2 composition is contraindicated in patients with known hypersensitivity to Zace2.

### 13. Therapeutic Uses of Zace2 Nucleotide Sequences

The present invention includes the use of Zace2 nucleotide sequences to provide Zace2 to a subject in need of such treatment. In addition, a therapeutic expression vector can be provided that inhibits Zace2 gene expression, such as an anti-sense molecule, a ribozyme, or an external guide sequence molecule. Although murine Zace2 nucleotide sequences can be used for these methods, compositions comprising human Zace2 nucleotide sequences are preferred for treatment of human subjects.

There are numerous approaches to introduce a Zace2 gene to a subject, including the use of recombinant host cells that express Zace2, delivery of naked

nucleic acid encoding *Zace2*, use of a cationic lipid carrier with a nucleic acid molecule that encodes *Zace2*, and the use of viruses that express *Zace2*, such as recombinant retroviruses, recombinant adeno-associated viruses, recombinant adenoviruses, and recombinant Herpes simplex viruses (see, for example, Mulligan, *Science* 260:926  
5 (1993), Rosenberg *et al.*, *Science* 242:1575 (1988), LaSalle *et al.*, *Science* 259:988 (1993), Wolff *et al.*, *Science* 247:1465 (1990), Breakfield and Deluca, *The New Biologist* 3:203 (1991)). In an *ex vivo* approach, for example, cells are isolated from a subject, transfected with a vector that expresses a *Zace2* gene, and then transplanted into the subject.

10 In order to effect expression of a *Zace2* gene, an expression vector is constructed in which a nucleotide sequence encoding a *Zace2* gene is operably linked to a core promoter, and optionally a regulatory element, to control gene transcription. The general requirements of an expression vector are described above.

Alternatively, a *Zace2* gene can be delivered using recombinant viral  
15 vectors, including for example, adenoviral vectors (e.g., Kass-Eisler *et al.*, *Proc. Nat'l Acad. Sci. USA* 90:11498 (1993), Kolls *et al.*, *Proc. Nat'l Acad. Sci. USA* 91:215 (1994), Li *et al.*, *Hum. Gene Ther.* 4:403 (1993), Vincent *et al.*, *Nat. Genet.* 5:130 (1993), and Zabner *et al.*, *Cell* 75:207 (1993)), adenovirus-associated viral vectors (Flotte *et al.*, *Proc. Nat'l Acad. Sci. USA* 90:10613 (1993)), alphaviruses such as  
20 Semliki Forest Virus and Sindbis Virus (Hertz and Huang, *J. Vir.* 66:857 (1992), Raju and Huang, *J. Vir.* 65:2501 (1991), and Xiong *et al.*, *Science* 243:1188 (1989)), herpes viral vectors (e.g., U.S. Patent Nos. 4,769,331, 4,859,587, 5,288,641 and 5,328,688), parvovirus vectors (Koering *et al.*, *Hum. Gene Therap.* 5:457 (1994)), pox virus vectors (Ozaki *et al.*, *Biochem. Biophys. Res. Comm.* 193:653 (1993), Panicali and Paoletti,  
25 *Proc. Nat'l Acad. Sci. USA* 79:4927 (1982)), pox viruses, such as canary pox virus or vaccinia virus (Fisher-Hoch *et al.*, *Proc. Nat'l Acad. Sci. USA* 86:317 (1989), and Flexner *et al.*, *Ann. N.Y. Acad. Sci.* 569:86 (1989)), and retroviruses (e.g., Baba *et al.*, *J. Neurosurg* 79:729 (1993), Ram *et al.*, *Cancer Res.* 53:83 (1993), Takamiya *et al.*, *J. Neurosci. Res* 33:493 (1992), Vile and Hart, *Cancer Res.* 53:962 (1993), Vile and Hart,  
30 *Cancer Res.* 53:3860 (1993), and Anderson *et al.*, U.S. Patent No. 5,399,346). Within various embodiments, either the viral vector itself, or a viral particle, which contains the viral vector may be utilized in the methods and compositions described below.

As an illustration of one system, adenovirus, a double-stranded DNA virus, is a well-characterized gene transfer vector for delivery of a heterologous nucleic  
35 acid molecule (for a review, see Becker *et al.*, *Meth. Cell Biol.* 43:161 (1994); Douglas and Curiel, *Science & Medicine* 4:44 (1997)). The adenovirus system offers several advantages including: (i) the ability to accommodate relatively large DNA inserts, (ii)

the ability to be grown to high-titer, (iii) the ability to infect a broad range of mammalian cell types, and (iv) the ability to be used with many different promoters including ubiquitous, tissue specific, and regulatable promoters. In addition, adenoviruses can be administered by intravenous injection, because the viruses are stable in the bloodstream.

Using adenovirus vectors where portions of the adenovirus genome are deleted, inserts are incorporated into the viral DNA by direct ligation or by homologous recombination with a co-transfected plasmid. In an exemplary system, the essential E1 gene is deleted from the viral vector, and the virus will not replicate unless the E1 gene is provided by the host cell. When intravenously administered to intact animals, adenovirus primarily targets the liver. Although an adenoviral delivery system with an E1 gene deletion cannot replicate in the host cells, the host's tissue will express and process an encoded heterologous protein. Host cells will also secrete the heterologous protein if the corresponding gene includes a secretory signal sequence. Secreted proteins will enter the circulation from tissue that expresses the heterologous gene (e.g., the highly vascularized liver).

Moreover, adenoviral vectors containing various deletions of viral genes can be used to reduce or eliminate immune responses to the vector. Such adenoviruses are E1-deleted, and in addition, contain deletions of E2A or E4 (Lusky *et al.*, *J. Virol.* 72:2022 (1998); Raper *et al.*, *Human Gene Therapy* 9:671 (1998)). The deletion of E2b has also been reported to reduce immune responses (Amalfitano *et al.*, *J. Virol.* 72:926 (1998)). By deleting the entire adenovirus genome, very large inserts of heterologous DNA can be accommodated. Generation of so called "gutless" adenoviruses, where all viral genes are deleted, are particularly advantageous for insertion of large inserts of heterologous DNA (for a review, see Yeh. and Perricaudet, *FASEB J.* 11:615 (1997)).

High titer stocks of recombinant viruses capable of expressing a therapeutic gene can be obtained from infected mammalian cells using standard methods. For example, recombinant herpes simplex virus can be prepared in Vero cells, as described by Brandt *et al.*, *J. Gen. Virol.* 72:2043 (1991), Herold *et al.*, *J. Gen. Virol.* 75:1211 (1994), Visalli and Brandt, *Virology* 185:419 (1991), Grau *et al.*, *Invest. Ophthalmol. Vis. Sci.* 30:2474 (1989), Brandt *et al.*, *J. Virol. Meth.* 36:209 (1992), and by Brown and MacLean (eds.), *HSV Virus Protocols* (Humana Press 1997).

Alternatively, an expression vector comprising a *Zace2* gene can be introduced into a subject's cells by lipofection *in vivo* using liposomes. Synthetic cationic lipids can be used to prepare liposomes for *in vivo* transfection of a gene encoding a marker (Felgner *et al.*, *Proc. Nat'l Acad. Sci. USA* 84:7413 (1987); Mackey *et al.*, *Proc. Nat'l Acad. Sci. USA* 85:8027 (1988)). The use of lipofection to introduce

exogenous genes into specific organs *in vivo* has certain practical advantages. Liposomes can be used to direct transfection to particular cell types, which is particularly advantageous in a tissue with cellular heterogeneity, such as the pancreas, liver, kidney, and brain. Lipids may be chemically coupled to other molecules for the purpose of targeting. Targeted peptides (*e.g.*, hormones or neurotransmitters), proteins such as antibodies, or non-peptide molecules can be coupled to liposomes chemically.

Electroporation is another alternative mode of administration. For example, Aihara and Miyazaki, *Nature Biotechnology* 16:867 (1998), have demonstrated the use of *in vivo* electroporation for gene transfer into muscle.

In an alternative approach to gene therapy, a therapeutic gene may encode a *Zace2* anti-sense RNA that inhibits the expression of *Zace2*. Suitable sequences for anti-sense molecules can be derived from the nucleotide sequences of *Zace2* disclosed herein.

Alternatively, an expression vector can be constructed in which a regulatory element is operably linked to a nucleotide sequence that encodes a ribozyme. Ribozymes can be designed to express endonuclease activity that is directed to a certain target sequence in a mRNA molecule (see, for example, Draper and Macejak, U.S. Patent No. 5,496,698, McSwiggen, U.S. Patent No. 5,525,468, Chowrira and McSwiggen, U.S. Patent No. 5,631,359, and Robertson and Goldberg, U.S. Patent No. 5,225,337). In the context of the present invention, ribozymes include nucleotide sequences that bind with *Zace2* mRNA.

In another approach, expression vectors can be constructed in which a regulatory element directs the production of RNA transcripts capable of promoting RNase P-mediated cleavage of mRNA molecules that encode a *Zace2* gene. According to this approach, an external guide sequence can be constructed for directing the endogenous ribozyme, RNase P, to a particular species of intracellular mRNA, which is subsequently cleaved by the cellular ribozyme (see, for example, Altman *et al.*, U.S. Patent No. 5,168,053, Yuan *et al.*, *Science* 263:1269 (1994), Pace *et al.*, international publication No. WO 96/18733, George *et al.*, international publication No. WO 96/21731, and Werner *et al.*, international publication No. WO 97/33991). Preferably, the external guide sequence comprises a ten to fifteen nucleotide sequence complementary to *Zace2* mRNA, and a 3'-NCCA nucleotide sequence, wherein N is preferably a purine. The external guide sequence transcripts bind to the targeted mRNA species by the formation of base pairs between the mRNA and the complementary external guide sequences, thus promoting cleavage of mRNA by RNase P at the nucleotide located at the 5'-side of the base-paired region.

In general, the dosage of a composition comprising a therapeutic vector having a *Zace2* nucleotide acid sequence, such as a recombinant virus, will vary depending upon such factors as the subject's age, weight, height, sex, general medical condition and previous medical history. Suitable routes of administration of therapeutic vectors include intravenous injection, intraarterial injection, intraperitoneal injection, intramuscular injection, intratumoral injection, and injection into a cavity that contains a tumor. As an illustration, Horton *et al.*, *Proc. Nat'l Acad. Sci. USA* 96:1553 (1999), demonstrated that intramuscular injection of plasmid DNA encoding interferon- $\alpha$  produces potent antitumor effects on primary and metastatic tumors in a murine model.

A composition comprising viral vectors, non-viral vectors, or a combination of viral and non-viral vectors of the present invention can be formulated according to known methods to prepare pharmaceutically useful compositions, whereby vectors or viruses are combined in a mixture with a pharmaceutically acceptable carrier. As noted above, a composition, such as phosphate-buffered saline is said to be a "pharmaceutically acceptable carrier" if its administration can be tolerated by a recipient subject. Other suitable carriers are well-known to those in the art (see, for example, *Remington's Pharmaceutical Sciences*, 19th Ed. (Mack Publishing Co. 1995), and *Gilman's the Pharmacological Basis of Therapeutics*, 7th Ed. (MacMillan Publishing Co. 1985)).

For purposes of therapy, a therapeutic gene expression vector, or a recombinant virus comprising such a vector, and a pharmaceutically acceptable carrier are administered to a subject in a therapeutically effective amount. A combination of an expression vector (or virus) and a pharmaceutically acceptable carrier is said to be administered in a "therapeutically effective amount" if the amount administered is physiologically significant. An agent is physiologically significant if its presence results in a detectable change in the physiology of a recipient subject. For example, common symptoms of Crohn's disease include chronic diarrhea with abdominal pain, fever, anorexia, weight loss, and a right lower quadrant mass. An agent used to treat Crohn's disease is physiologically significant if its presence alleviates at least one of these symptoms.

When the subject treated with a therapeutic gene expression vector or a recombinant virus is a human, then the therapy is preferably somatic cell gene therapy. That is, the preferred treatment of a human with a therapeutic gene expression vector or a recombinant virus does not entail introducing into cells a nucleic acid molecule that can form part of a human germ line and be passed onto successive generations (*i.e.*, human germ line gene therapy).

#### 14. Production of Transgenic Mice

Transgenic mice can be engineered to over-express the *Zace2* gene in all tissues or under the control of a tissue-specific or tissue-preferred regulatory element. These over-producers of *Zace2* can be used to characterize the phenotype that results from over-expression, and the transgenic animals can serve as models for human disease caused by excess *Zace2*. Transgenic mice that over-express *Zace2* also provide model bioreactors for production of *Zace2* in the milk or blood of larger animals. Methods for producing transgenic mice are well-known to those of skill in the art (see, for example, Jacob, "Expression and Knockout of Interferons in Transgenic Mice," in *Overexpression and Knockout of Cytokines in Transgenic Mice*, Jacob (ed.), pages 111-124 (Academic Press, Ltd. 1994), Monastersky and Robl (eds.), *Strategies in Transgenic Animal Science* (ASM Press 1995), and Abbud and Nilson, "Recombinant Protein Expression in Transgenic Mice," in *Gene Expression Systems: Using Nature for the Art of Expression*, Fernandez and Hoeffler (eds.), pages 367-397 (Academic Press, Inc. 1999)).

For example, a method for producing a transgenic mouse that expresses a *Zace2* gene can begin with adult, fertile males (studs) (B6C3f1, 2-8 months of age (Taconic Farms, Germantown, NY)), vasectomized males (duds) (B6D2f1, 2-8 months, (Taconic Farms)), prepubescent fertile females (donors) (B6C3f1, 4-5 weeks, (Taconic Farms)) and adult fertile females (recipients) (B6D2f1, 2-4 months, (Taconic Farms)). The donors are acclimated for one week and then injected with approximately 8 IU/mouse of Pregnant Mare's Serum gonadotrophin (Sigma Chemical Company; St. Louis, MO) I.P., and 46-47 hours later, 8 IU/mouse of human Chorionic Gonadotropin (hCG (Sigma)) I.P. to induce superovulation. Donors are mated with studs subsequent to hormone injections. Ovulation generally occurs within 13 hours of hCG injection. Copulation is confirmed by the presence of a vaginal plug the morning following mating.

Fertilized eggs are collected under a surgical scope. The oviducts are collected and eggs are released into urinalysis slides containing hyaluronidase (Sigma). Eggs are washed once in hyaluronidase, and twice in Whitten's W640 medium (described, for example, by Menino and O'Claray, *Biol. Reprod.* 77:159 (1986), and Dienhart and Downs, *Zygote* 4:129 (1996)) that has been incubated with 5% CO<sub>2</sub>, 5% O<sub>2</sub>, and 90% N<sub>2</sub> at 37°C. The eggs are then stored in a 37°C/5% CO<sub>2</sub> incubator until microinjection.

Ten to twenty micrograms of plasmid DNA containing a *Zace2* encoding sequence is linearized, gel-purified, and resuspended in 10 mM Tris-HCl (pH 7.4), 0.25 mM EDTA (pH 8.0), at a final concentration of 5-10 nanograms per

microliter for microinjection. For example, the *Zace2* encoding sequences can encode a polypeptide comprising amino acid residues 19 to 738 of SEQ ID NOs:2, 6, or 9.

5 Plasmid DNA is microinjected into harvested eggs contained in a drop of W640 medium overlaid by warm, CO<sub>2</sub>-equilibrated mineral oil. The DNA is drawn into an injection needle (pulled from a 0.75mm ID, 1mm OD borosilicate glass capillary), and injected into individual eggs. Each egg is penetrated with the injection needle, into one or both of the haploid pronuclei.

10 Picoliters of DNA are injected into the pronuclei, and the injection needle withdrawn without coming into contact with the nucleoli. The procedure is repeated until all the eggs are injected. Successfully microinjected eggs are transferred into an organ tissue-culture dish with pre-gassed W640 medium for storage overnight in a 37°C/5% CO<sub>2</sub> incubator.

The following day, two-cell embryos are transferred into pseudopregnant recipients. The recipients are identified by the presence of copulation plugs, after copulating with vasectomized duds. Recipients are anesthetized and shaved  
15 on the dorsal left side and transferred to a surgical microscope. A small incision is made in the skin and through the muscle wall in the middle of the abdominal area outlined by the ribcage, the saddle, and the hind leg, midway between knee and spleen. The reproductive organs are exteriorized onto a small surgical drape. The fat pad is stretched out over the surgical drape, and a baby serrefine (Roboz, Rockville, MD) is attached to the fat pad and left hanging over the back of the mouse, preventing the  
20 organs from sliding back in.

With a fine transfer pipette containing mineral oil followed by alternating W640 and air bubbles, 12-17 healthy two-cell embryos from the previous  
25 day's injection are transferred into the recipient. The swollen ampulla is located and holding the oviduct between the ampulla and the bursa, a nick in the oviduct is made with a 28 g needle close to the bursa, making sure not to tear the ampulla or the bursa.

The pipette is transferred into the nick in the oviduct, and the embryos are blown in, allowing the first air bubble to escape the pipette. The fat pad is gently  
30 pushed into the peritoneum, and the reproductive organs allowed to slide in. The peritoneal wall is closed with one suture and the skin closed with a wound clip. The mice recuperate on a 37°C slide warmer for a minimum of four hours.

35 TL<sub>2</sub> recipients are returned to cages in pairs, and allowed 19-21 days gestation. After birth, 19-21 days postpartum is allowed before weaning. The weanlings are sexed and placed into separate sex cages, and a 0.5 cm biopsy (used for genotyping) is snipped off the tail with clean scissors.

Genomic DNA is prepared from the tail snips using, for example, a QIAGEN DNEASY kit following the manufacturer's instructions. Genomic DNA is analyzed by PCR using primers designed to amplify a *Zace2* gene or a selectable marker gene that was introduced in the same plasmid. After animals are confirmed to be transgenic, they are back-crossed into an inbred strain by placing a transgenic female with a wild-type male, or a transgenic male with one or two wild-type female(s). As pups are born and weaned, the sexes are separated, and their tails snipped for genotyping.

To check for expression of a transgene in a live animal, a partial hepatectomy is performed. A surgical prep is made of the upper abdomen directly below the zyphoid process. Using sterile technique, a small 1.5-2 cm incision is made below the sternum and the left lateral lobe of the liver exteriorized. Using 4-0 silk, a tie is made around the lower lobe securing it outside the body cavity. An atraumatic clamp is used to hold the tie while a second loop of absorbable Dexon (American Cyanamid; Wayne, N.J.) is placed proximal to the first tie. A distal cut is made from the Dexon tie and approximately 100 mg of the excised liver tissue is placed in a sterile petri dish. The excised liver section is transferred to a 14 ml polypropylene round bottom tube and snap frozen in liquid nitrogen and then stored on dry ice. The surgical site is closed with suture and wound clips, and the animal's cage placed on a 37°C heating pad for 24 hours post operatively. The animal is checked daily post operatively and the wound clips removed 7-10 days after surgery. The expression level of *Zace2* mRNA is examined for each transgenic mouse using an RNA solution hybridization assay or polymerase chain reaction.

In addition to producing transgenic mice that over-express *Zace2*, it is useful to engineer transgenic mice with either abnormally low or no expression of the gene. Such transgenic mice provide useful models for diseases associated with a lack of *Zace2*. As discussed above, *Zace2* gene expression can be inhibited using anti-sense genes, ribozyme genes, or external guide sequence genes. To produce transgenic mice that under-express the *Zace2* gene, such inhibitory sequences are targeted to *Zace2* mRNA. Methods for producing transgenic mice that have abnormally low expression of a particular gene are known to those in the art (see, for example, Wu *et al.*, "Gene Underexpression in Cultured Cells and Animals by Antisense DNA and RNA Strategies," in *Methods in Gene Biotechnology*, pages 205-224 (CRC Press 1997)).

An alternative approach to producing transgenic mice that have little or no *Zace2* gene expression is to generate mice having at least one normal *Zace2* allele replaced by a nonfunctional *Zace2* gene. One method of designing a nonfunctional *Zace2* gene is to insert another gene, such as a selectable marker gene, within a nucleic



acid molecule that encodes Zace2. Standard methods for producing these so-called "knockout mice" are known to those skilled in the art (see, for example, Jacob, "Expression and Knockout of Interferons in Transgenic Mice," in *Overexpression and Knockout of Cytokines in Transgenic Mice*, Jacob (ed.), pages 111-124 (Academic Press, Ltd. 1994), and Wu *et al.*, "New Strategies for Gene Knockout," in *Methods in Gene Biotechnology*, pages 339-365 (CRC Press 1997)).

The present invention, thus generally described, will be understood more readily by reference to the following examples, which are provided by way of illustration and are not intended to be limiting of the present invention.

#### EXAMPLE 1

##### *Expression of the Human Zace2 Gene*

Northern analyses were performed using Human Multiple Tissue Blots (CLONTECH Laboratories, Inc., Palo Alto, CA). A human Zace2 cDNA probe comprising a 5' portion of the nucleotide sequence of SEQ ID NO:1 was radioactively labeled using the Rediprime II labeling kit (Amersham Pharmacia Biotech, Inc.; Piscataway, NJ) according to the manufacturer's protocol. The probe was purified using a NUCTRAP push column (STRATAGENE; La Jolla, CA). EXPRESSHYB (CLONTECH) solution was used for the prehybridization and hybridization solutions for the northern blots. Hybridization took place overnight at 65°C. Following hybridization, the blots were washed four times at 25°C in 2xSSC with 0.05% SDS at room temperature, and then, twice at 50°C in 0.1xSSC with 0.1% SDS. The results showed that the human Zace2 gene is predominantly expressed as a mRNA species of about four kilobases by testicular tissue, and that there is less expression in kidney, thyroid, small intestine, colon, heart, and potentially, adrenal tissues. In contrast, little or no expression was observed in spleen, thymus, prostate, and ovarian tissues.

#### EXAMPLE 2

##### *Expression of the Murine Zace2 Gene*

Analyses of murine Zace2 gene expression were performed using Mouse Multiple Tissue Blots from OriGene Technologies, Inc. (Rockville, MD) and CLONTECH Laboratories, Inc. (Palo Alto, CA). A DNA probe of about 500 base pairs was prepared using a 35-cycle polymerase chain reaction with EX TAQ (PANVERA Corporation; Madison, WI) and an annealing temperature of 60°C. Oligo 21982 (5'-GACTCCGATCATCAAGCGTCAACTA-3'; SEQ ID NO:10) was used to generate the 5'-end of the DNA probe, and the 3'-end was generated with oligo 22997 (5'-

GGCAGGGAGGCATCCAGTGG-3'; SEQ ID NO:11). The DNA probe was gel purified using QIAquick gel extraction kit (QIAGEN, Inc.; Valencia, CA), radioactively labeled with  $^{32}\text{P}$  using the Rediprime II DNA labeling system (Amersham Pharmacia Biotech, Inc.; Piscataway, NJ) according to the manufacturer's specifications, and the radiolabeled probe was purified using a NUCTRAP push column (STRATAGENE, La Jolla, CA). EXPRESSHYB (CLONTECH) solution was used for prehybridization and hybridization. Following an overnight hybridization at 65°C, the blots were washed four times at 25°C in 2xSSC with 0.05% SDS at room temperature, and then, twice at 50°C in 0.1xSSC with 0.1% SDS. *Zace2* transcripts of about 3.5 kilobases and about 4.0 kilobases were observed in the following murine tissues: 17-day embryo, kidney, small intestine, and skin. Hybridization with a dot blot also produced signals in kidney, smooth muscle (small intestine), and 17-day embryo tissues. RNA from small intestine tissue produced the strongest signal.

## CLAIMS

What is claimed is:

1. An isolated polypeptide, comprising an amino acid sequence that is at least 70% identical to a reference amino acid sequence selected from the group consisting of: (a) amino acid residues 19 to 805 of either SEQ ID NO:6 or SEQ ID NO:9, (b) amino acid residues 19 to 738 of either SEQ ID NO:6 or SEQ ID NO:9, (c) amino acid residues 19 to 708 of either SEQ ID NO:6 or SEQ ID NO:9, (d) amino acid residues 19 to 613 of either SEQ ID NO:6 or SEQ ID NO:9, (e) amino acid residues 133 to 542 of either SEQ ID NO:6 or SEQ ID NO:9, (f) amino acid residues 344 to 542 of either SEQ ID NO:6 or SEQ ID NO:9, and (g) amino acid residues 371 to 402 of SEQ ID NO:6,

wherein the isolated polypeptide either (a) specifically binds with an antibody that specifically binds with a polypeptide consisting of the amino acid sequence of either SEQ ID NO:6 or SEQ ID NO:9, or (b) exhibits dipeptidyl carboxypeptidase activity.

2. The isolated polypeptide of claim 1, wherein the isolated polypeptide has an amino acid sequence that is at least 80% identical to the reference amino acid sequence.

3. The isolated polypeptide of claim 1, wherein the isolated polypeptide has an amino acid sequence that is at least 90% identical to the reference amino acid sequence.

4. The isolated polypeptide of claim 1, wherein the polypeptide is a metallopeptidase.

5. The isolated polypeptide of claim 1, wherein the isolated polypeptide comprises amino acid residues 19 to 738 of either SEQ ID NO:6 or SEQ ID NO:9.

6. The isolated polypeptide of claim 1, wherein the polypeptide comprises an amino acid sequence comprising the motif "[GSTALIVN]-x-x-H-E-[LIVMFYW]-{DEHRKP}-H-x-[LIVMFYWGSPQ]," where "x" is any amino acid residue, acceptable amino acid residues are listed between square brackets, and unacceptable amino acid residues are listed between braces.

7. The isolated polypeptide of claim 1, comprising the amino acid sequence of either SEQ ID NO:6 or SEQ ID NO:9.

8. A variant Zace2 polypeptide, wherein the amino acid sequence of the variant polypeptide shares an identity with the amino acid sequence of either SEQ ID NO:6 or SEQ ID NO:9 selected from the group consisting of at least 70% identity, at least 80% identity, at least 90% identity, at least 95% identity, or greater than 95% identity, and wherein any difference between the amino acid sequence of the variant polypeptide and the amino acid sequence of either SEQ ID NO:6 or SEQ ID NO:9 is due to one or more conservative amino acid substitutions.

9. A variant Zace2 polypeptide, wherein the amino acid sequence of the variant polypeptide is a mutation of the amino acid sequence of either SEQ ID NO:6 or SEQ ID NO:9, which includes at least one amino acid substitution selected from the group consisting of: Gln<sup>24</sup>, Val<sup>59</sup>, Gly<sup>66</sup>, Tyr<sup>83</sup>, Val<sup>93</sup>, Val<sup>107</sup>, Val<sup>209</sup>, Val<sup>212</sup>, His<sup>228</sup>, Gly<sup>286</sup>, Asp<sup>335</sup>, Gly<sup>337</sup>, Val<sup>339</sup>, Ala<sup>342</sup>, Lys<sup>353</sup>, Lys<sup>465</sup>, Val<sup>491</sup>, Leu<sup>520</sup>, Lys<sup>577</sup>, Arg<sup>582</sup>, Lys<sup>600</sup>, Tyr<sup>641</sup>, Val<sup>658</sup>, Ile<sup>663</sup>, Gly<sup>666</sup>, Ile<sup>679</sup>, Ile<sup>694</sup>, Thr<sup>698</sup>, Ala<sup>714</sup>, Ser<sup>740</sup>, Val<sup>745</sup>, Gly<sup>751</sup>, Val<sup>752</sup>, Ile<sup>753</sup>, Arg<sup>766</sup>, Arg<sup>775</sup>, Thr<sup>798</sup>, and Val<sup>801</sup>.

10. An isolated nucleic acid molecule that encodes a Zace2 polypeptide, wherein the nucleic acid molecule is selected from the group consisting of (a) a nucleic acid molecule encoding an amino acid sequence that comprises amino acid residues 19 to 738 of SEQ ID NO:6, (b) a nucleic acid molecule encoding an amino acid sequence that comprises amino acid residues 19 to 738 of SEQ ID NO:9, and (c) a nucleic acid molecule that remains hybridized following stringent wash conditions to a nucleic acid molecule comprising the nucleotide sequence of nucleotides 106-2520 of SEQ ID NO:5, or the complement of nucleotides 106-2520 of SEQ ID NO:5.

11. An expression vector, comprising the isolated nucleic acid molecule of claim 10, wherein the nucleic acid molecule encodes an amino acid sequence comprising amino acid residues 19 to 738 of SEQ ID NO:6 or amino acid residues 19 to 738 of SEQ ID NO:9, a transcription promoter, and a transcription terminator, wherein the promoter is operably linked with the nucleic acid molecule, and wherein the nucleic acid molecule is operably linked with the transcription terminator.

12. A recombinant host cell comprising the expression vector of claim 11, wherein the host cell is selected from the group consisting of bacterium, yeast cell, fungal cell, insect cell, avian cell, mammalian cell, and plant cell.

13. A method of using the expression vector of claim 11 to produce *Zace2* protein, comprising culturing recombinant host cells that comprise the expression vector and that produce the *Zace2* protein.

14. The method of claim 13, further comprising isolating the *Zace2* protein from the cultured recombinant host cells.

15. An antibody or antibody fragment that specifically binds with the polypeptide of claim 5.

16. An anti-idiotypic antibody, or anti-idiotypic antibody fragment, that specifically binds with the antibody or antibody fragment of claim 15, wherein the anti-idiotypic antibody, or anti-idiotypic antibody fragment, possesses dipeptidyl carboxypeptidase activity.

17. A method of detecting the presence of *Zace2* RNA in a biological sample, comprising :

(a) contacting a *Zace2* nucleic acid probe under hybridizing conditions with either (i) test RNA molecules isolated from the biological sample, or (ii) nucleic acid molecules synthesized from the isolated RNA molecules, wherein the probe has a nucleotide sequence comprising a portion of the nucleotide sequence of SEQ ID NO:5, or the complement of the nucleotide sequence of SEQ ID NO:5, and

(b) detecting the formation of hybrids of the nucleic acid probe and either the test RNA molecules or the synthesized nucleic acid molecules,

wherein the presence of the hybrids indicates the presence of *Zace2* RNA in the biological sample.

18. A method of detecting the presence of *Zace2* in a biological sample, comprising the steps of:

(a) contacting the biological sample with an antibody, or an antibody fragment, that specifically binds with a polypeptide consisting of the amino acid sequence of either SEQ ID NO:6 or SEQ ID NO:9, wherein the contacting is

performed under conditions that allow the binding of the antibody or antibody fragment to the biological sample, and

(b) detecting any of the bound antibody or bound antibody fragment.

19. A recombinant virus, comprising the expression vector of claim 11.

20. A composition, comprising a carrier and at least one of the expression vector of claim 11, or a recombinant virus that comprises the expression vector of claim 11.

21. A composition, comprising the isolated polypeptide of claim 5 and a carrier.

## SEQUENCE LISTING

&lt;110&gt; ZymoGenetics, Inc.

&lt;120&gt; ZACE2: A HUMAN METALLOENZYME

&lt;130&gt; 99-24PC

&lt;160&gt; 11

&lt;170&gt; FastSEQ for Windows Version 3.0

&lt;210&gt; 1

&lt;211&gt; 3334

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (35)...(2449)

&lt;400&gt; 1

attcagtgga tgtgatcttg gctcacaggg gacg atg tca agc tct tcc tgg ctc	55
Met Ser Ser Ser Ser Trp Leu	
1 5	

ctt ctc agc ctt gtt gct gta act gct gct cag tcc acc att gag gaa	103
Leu Leu Ser Leu Val Ala Val Thr Ala Ala Gln Ser Thr Ile Glu Glu	
10 15 20	

cag gcc aag aca ttt ttg gac aag ttt aac cac gaa gcc gaa gac ctg	151
Gln Ala Lys Thr Phe Leu Asp Lys Phe Asn His Glu Ala Glu Asp Leu	
25 30 35	

ttc tat caa agt tca ctt gct tct tgg aat tat aac acc aat att act	199
Phe Tyr Gln Ser Ser Leu Ala Ser Trp Asn Tyr Asn Thr Asn Ile Thr	
40 45 50 55	

gaa gag aat gtc caa aac atg aat aat gct ggg gac aaa tgg tct gcc	247
Glu Glu Asn Val Gln Asn Met Asn Asn Ala Gly Asp Lys Trp Ser Ala	
60 65 70	

ttt tta aag gaa cag tcc aca ctt gcc caa atg tat cca cta caa gaa	295
Phe Leu Lys Glu Gln Ser Thr Leu Ala Gln Met Tyr Pro Leu Gln Glu	

75	80	85	
att cag aat ctc aca gtc aag ctt cag ctg cag gct ctt cag caa aat			343
Ile Gln Asn Leu Thr Val Lys Leu Gln Leu Gln Ala Leu Gln Gln Asn			
90	95	100	
ggg tct tca gtg ctc tca gaa gac aag agc aaa cgg ttg aac aca att			391
Gly Ser Ser Val Leu Ser Glu Asp Lys Ser Lys Arg Leu Asn Thr Ile			
105	110	115	
cta aat aca atg agc acc atc tac agt act gga aaa gtt tgt aac cca			439
Leu Asn Thr Met Ser Thr Ile Tyr Ser Thr Gly Lys Val Cys Asn Pro			
120	125	130	135
gat aat cca caa gaa tgc tta tta ctt gaa cca ggt ttg aat gaa ata			487
Asp Asn Pro Gln Glu Cys Leu Leu Leu Glu Pro Gly Leu Asn Glu Ile			
140	145	150	
atg gca aac agt tta gac tac aat gag agg ctc tgg gct tgg gaa agc			535
Met Ala Asn Ser Leu Asp Tyr Asn Glu Arg Leu Trp Ala Trp Glu Ser			
155	160	165	
tgg aga tct gag gtc ggc aag cag ctg agg cca tta tat gaa gag tat			583
Trp Arg Ser Glu Val Gly Lys Gln Leu Arg Pro Leu Tyr Glu Glu Tyr			
170	175	180	
gtg gtc ttg aaa aat gag atg gca aga gca aat cat tat gag gac tat			631
Val Val Leu Lys Asn Glu Met Ala Arg Ala Asn His Tyr Glu Asp Tyr			
185	190	195	
ggg gat tat tgg aga gga gac tat gaa gta aat ggg gta gat ggc tat			679
Gly Asp Tyr Trp Arg Gly Asp Tyr Glu Val Asn Gly Val Asp Gly Tyr			
200	205	210	215
gac tac agc cgc ggc cag ttg att gaa gat gtg gaa cat acc ttt gaa			727
Asp Tyr Ser Arg Gly Gln Leu Ile Glu Asp Val Glu His Thr Phe Glu			
220	225	230	
gag att aaa cca tt. tat gaa cat ctt cat gcc tat gtg agg gca aag			775
Glu Ile Lys Pro Leu Tyr Glu His Leu His Ala Tyr Val Arg Ala Lys			
235	240	245	
ttg atg aat gcc tat cct tcc tat atc agt cca att gga tgc ctc cct			823
Leu Met Asn Ala Tyr Pro Ser Tyr Ile Ser Pro Ile Gly Cys Leu Pro			
250	255	260	



gct cat ttg ctt ggt gat atg tgg ggt aga ttt tgg aca aat ctg tac	871
Ala His Leu Leu Gly Asp Met Trp Gly Arg Phe Trp Thr Asn Leu Tyr	
265 270 275	
tct ttg aca gtt ccc ttt gga cag aaa cca aac ata gat gtt act gat	919
Ser Leu Thr Val Pro Phe Gly Gln Lys Pro Asn Ile Asp Val Thr Asp	
280 285 290 295	
gca atg gtg gac cag gcc tgg gat gca cag aga ata ttc aag gag gcc	967
Ala Met Val Asp Gln Ala Trp Asp Ala Gln Arg Ile Phe Lys Glu Ala	
300 305 310	
gag aag ttc ttt gta tct gtt ggt ctt cct aat atg act caa gga ttc	1015
Glu Lys Phe Phe Val Ser Val Gly Leu Pro Asn Met Thr Gln Gly Phe	
315 320 325	
tgg gaa aat tcc atg cta acg gac cca gga aat gtt cag aaa gca gtc	1063
Trp Glu Asn Ser Met Leu Thr Asp Pro Gly Asn Val Gln Lys Ala Val	
330 335 340	
tgc cat ccc aca gct tgg gac ctg ggg aag ggc gac ttc agg atc ctt	1111
Cys His Pro Thr Ala Trp Asp Leu Gly Lys Gly Asp Phe Arg Ile Leu	
345 350 355	
atg tgc aca aag gtg aca atg gac gac ttc ctg aca gct cat cat gag	1159
Met Cys Thr Lys Val Thr Met Asp Asp Phe Leu Thr Ala His His Glu	
360 365 370 375	
atg ggg cat atc cag tat gat atg gca tat gct gca caa cct ttt ctg	1207
Met Gly His Ile Gln Tyr Asp Met Ala Tyr Ala Ala Gln Pro Phe Leu	
380 385 390	
cta aga aat gga gct aat gaa gga ttc cat gaa gct gtt ggg gaa atc	1255
Leu Arg Asn Gly Ala Asn Glu Gly Phe His Glu Ala Val Gly Glu Ile	
395 400 405	
atg tca ctt tct gca gcc aca cct aag cat tta aaa tcc att ggt ctt	1303
Met Ser Leu Ser Ala Ala Thr Pro Lys His Leu Lys Ser Ile Gly Leu	
410 415 420	
ctg tca ccc gat ttt caa gaa gac aat gaa aca gaa ata aac ttc ctg	1351
Leu Ser Pro Asp Phe Gln Glu Asp Asn Glu Thr Glu Ile Asn Phe Leu	
425 430 435	

ctc aaa caa gca ctc acg att gtt ggg act ctg cca ttt act tac atg Leu Lys Gln Ala Leu Thr Ile Val Gly Thr Leu Pro Phe Thr Tyr Met 440 445 450 455	1399
tta gag aag tgg agg tgg atg gtc ttt aaa ggg gaa att ccc aaa gac Leu Glu Lys Trp Arg Trp Met Val Phe Lys Gly Glu Ile Pro Lys Asp 460 465 470	1447
cag tgg atg aaa aag tgg tgg gag atg aag cga gag ata gtt ggg gtg Gln Trp Met Lys Lys Trp Trp Glu Met Lys Arg Glu Ile Val Gly Val 475 480 485	1495
gtg gaa cct gtg ccc cat gat gaa aca tac tgt gac ccc gca tct ctg Val Glu Pro Val Pro His Asp Glu Thr Tyr Cys Asp Pro Ala Ser Leu 490 495 500	1543
ttc cat gtt tct aat gat tac tca ttc att cga tat tac aca agg acc Phe His Val Ser Asn Asp Tyr Ser Phe Ile Arg Tyr Tyr Thr Arg Thr 505 510 515	1591
ctt tac caa ttc cag ttt caa gaa gca ctt tgt caa gca gct aaa cat Leu Tyr Gln Phe Gln Phe Gln Glu Ala Leu Cys Gln Ala Ala Lys His 520 525 530 535	1639
gaa ggc cct ctg cac aaa tgt gac atc tca aac tct aca gaa gct gga Glu Gly Pro Leu His Lys Cys Asp Ile Ser Asn Ser Thr Glu Ala Gly 540 545 550	1687
cag aaa ctg ttc aat atg ctg agg ctt gga aaa tca gaa ccc tgg acc Gln Lys Leu Phe Asn Met Leu Arg Leu Gly Lys Ser Glu Pro Trp Thr 555 560 565	1735
cta gca ttg gaa aat gtt gta gga gca aag aac atg aat gta agg cca Leu Ala Leu Glu Asn Val Val Gly Ala Lys Asn Met Asn Val Arg Pro 570 575 580	1783
ctg ctc aac tac ttt gag ccc tta ttt acc tgg ctg aaa gac cag aac Leu Leu Asn Tyr Phe Glu Pro Leu Phe Thr Trp Leu Lys Asp Gln Asn 585 590 595	1831
aag aat tct ttt gtg gga tgg agt acc gac tgg agt cca tat gca gac Lys Asn Ser Phe Val Gly Trp Ser Thr Asp Trp Ser Pro Tyr Ala Asp 600 605 610 615	1879

caa agc atc aaa gtg agg ata agc cta aaa tca gct ctt gga gat aaa 1927  
 Gln Ser Ile Lys Val Arg Ile Ser Leu Lys Ser Ala Leu Gly Asp Lys  
                   620                                  625                                  630

gca tat gaa tgg aac gac aat gaa atg tac ctg ttc cga tca tct gtt 1975  
 Ala Tyr Glu Trp Asn Asp Asn Glu Met Tyr Leu Phe Arg Ser Ser Val  
                   635                                  640                                  645

gca tat gct atg agg cag tac ttt tta aaa gta aaa aat cag atg att 2023  
 Ala Tyr Ala Met Arg Gln Tyr Phe Leu Lys Val Lys Asn Gln Met Ile  
                   650                                  655                                  660

ctt ttt ggg gag gag gat gtg cga gtg gct aat ttg aaa cca aga atc 2071  
 Leu Phe Gly Glu Glu Asp Val Arg Val Ala Asn Leu Lys Pro Arg Ile  
                   665                                  670                                  675

tcc ttt aat ttc ttt gtc act gca cct aaa aat gtg tct gat atc att 2119  
 Ser Phe Asn Phe Phe Val Thr Ala Pro Lys Asn Val Ser Asp Ile Ile  
                   680                                  685                                  690                                  695

cct aga act gaa gtt gaa aag gcc atc agg atg tcc cgg agc cgt atc 2167  
 Pro Arg Thr Glu Val Glu Lys Ala Ile Arg Met Ser Arg Ser Arg Ile  
                   700                                  705                                  710

aat gat gct ttc cgt ctg aat gac aac agc cta gag ttt ctg ggg ata 2215  
 Asn Asp Ala Phe Arg Leu Asn Asp Asn Ser Leu Glu Phe Leu Gly Ile  
                   715                                  720                                  725

cag cca aca ctt gga cct cct aac cag ccc cct gtt tcc ata tgg ctg 2263  
 Gln Pro Thr Leu Gly Pro Pro Asn Gln Pro Pro Val Ser Ile Trp Leu  
                   730                                  735                                  740

att gtt ttt gga gtt gtg atg gga gtg ata gtg gtt ggc att gtc atc 2311  
 Ile Val Phe Gly Val Val Met Gly Val Ile Val Val Gly Ile Val Ile  
                   745                                  750                                  755

ctg atc ttc act ggg atc aga gat cgg aag aag aaa aat aaa gca aga 2359  
 Leu Ile Phe Thr Gly Ile Arg Asp Arg Lys Lys Lys Asn Lys Ala Arg  
                   760                                  765                                  770                                  775

agt gga gaa aat cct tat gcc tcc atc gat att agc aaa gga gaa aat 2407  
 Ser Gly Glu Asn Pro Tyr Ala Ser Ile Asp Ile Ser Lys Gly Glu Asn  
                   780                                  785                                  790

aat cca gga ttc caa aac act gat gat gtt cag acc tcc ttt 2449  
 Asn Pro Gly Phe Gln Asn Thr Asp Asp Val Gln Thr Ser Phe  
 795 800 805

tagaaaaatc tatgtttttc ctcttgaggt gattttgttg tatgtaaatg ttaatttcat 2509  
 ggtatagaaa atataagatg ataaagatat cattaaatgt caaaactatg actctgttca 2569  
 gaaaaaaaaat tgtccaaaga caacatggcc aaggagagag catcttcatt gacattgctt 2629  
 tcagtattta tttctgtctc tggatttgac tttctgttctg tttcttaata aggattttgt 2689  
 attagagtat attagggaaa gtgtgtatatt ggtctcacag gctgttcagg gataatctaa 2749  
 atgtaaatgt ctgttgaatt tctgaagttg aaaacaagga tatatcattg gagcaagtgt 2809  
 tggatcttgt atggaatag gatggatcac ttgtaaggac agtgcctggg aactggtgta 2869  
 gctgcaagga ttgagaatgg catgcattag ctcaacttca tttaatccat tgtcaaggat 2929  
 gacatgcttt ctccacagta actcagttca agtactatgg tgatttgctt acagtgatgt 2989  
 ttggaatcga tcatgctttc ttcaaggtga caggtctaaa gagagaagaa tccaggggaa 3049  
 aggtagagga cattgctttt tcaacttcaa ggtgcttgat caacatctcc ctgacaacac 3109  
 aaaactagag ccaggggctt ccgtgaactc ccagagcatg cctgatagaa actcatttct 3169  
 actgttctct aactgtggag tgaatggaaa ttccaactgt atgttcaccc tctgaagtgg 3229  
 gtacccagtc tcttaaactc tttgtatttg ctacacagtgt ttgagcagtg ctgagcacia 3289  
 agcagacact caataaatgc tagattttaca cactccttgt gctta 3334

<210> 2

<211> 805

<212> PRT

<213> Homo sapiens

<400> 2

Met Ser Ser Ser Ser Trp Leu Leu Leu Ser Leu Val Ala Val Thr Ala  
 1 5 10 15  
 Ala Gln Ser Thr Ile Glu Glu Gln Ala Lys Thr Phe Leu Asp Lys Phe  
 20 25 30  
 Asn His Glu Ala Glu Asp Leu Phe Tyr Gln Ser Ser Leu Ala Ser Trp  
 35 40 45  
 Asn Tyr Asn Thr Asn Ile Thr Glu Glu Asn Val Gln Asn Met Asn Asn  
 50 55 60  
 Ala Gly Asp Lys Trp Ser Ala Phe Leu Lys Glu Gln Ser Thr Leu Ala  
 65 70 75 80  
 Gln Met Tyr Pro Leu Gln Glu Ile Gln Asn Leu Thr Val Lys Leu Gln  
 85 90 95  
 Leu Gln Ala Leu Gln Gln Asn Gly Ser Ser Val Leu Ser Glu Asp Lys  
 100 105 110  
 Ser Lys Arg Leu Asn Thr Ile Leu Asn Thr Met Ser Thr Ile Tyr Ser  
 115 120 125  
 Thr Gly Lys Val Cys Asn Pro Asp Asn Pro Gln Glu Cys Leu Leu Leu  
 130 135 140  
 Glu Pro Gly Leu Asn Glu Ile Met Ala Asn Ser Leu Asp Tyr Asn Glu

145                      150                      155                      160  
 Arg Leu Trp Ala Trp Glu Ser Trp Arg Ser Glu Val Gly Lys Gln Leu  
                                  165                      170                      175  
 Arg Pro Leu Tyr Glu Glu Tyr Val Val Leu Lys Asn Glu Met Ala Arg  
                                  180                      185                      190  
 Ala Asn His Tyr Glu Asp Tyr Gly Asp Tyr Trp Arg Gly Asp Tyr Glu  
                                  195                      200                      205  
 Val Asn Gly Val Asp Gly Tyr Asp Tyr Ser Arg Gly Gln Leu Ile Glu  
                                  210                      215                      220  
 Asp Val Glu His Thr Phe Glu Glu Ile Lys Pro Leu Tyr Glu His Leu  
 225                      230                      235                      240  
 His Ala Tyr Val Arg Ala Lys Leu Met Asn Ala Tyr Pro Ser Tyr Ile  
                                  245                      250                      255  
 Ser Pro Ile Gly Cys Leu Pro Ala His Leu Leu Gly Asp Met Trp Gly  
                                  260                      265                      270  
 Arg Phe Trp Thr Asn Leu Tyr Ser Leu Thr Val Pro Phe Gly Gln Lys  
                                  275                      280                      285  
 Pro Asn Ile Asp Val Thr Asp Ala Met Val Asp Gln Ala Trp Asp Ala  
                                  290                      295                      300  
 Gln Arg Ile Phe Lys Glu Ala Glu Lys Phe Phe Val Ser Val Gly Leu  
 305                      310                      315                      320  
 Pro Asn Met Thr Gln Gly Phe Trp Glu Asn Ser Met Leu Thr Asp Pro  
                                  325                      330                      335  
 Gly Asn Val Gln Lys Ala Val Cys His Pro Thr Ala Trp Asp Leu Gly  
                                  340                      345                      350  
 Lys Gly Asp Phe Arg Ile Leu Met Cys Thr Lys Val Thr Met Asp Asp  
                                  355                      360                      365  
 Phe Leu Thr Ala His His Glu Met Gly His Ile Gln Tyr Asp Met Ala  
                                  370                      375                      380  
 Tyr Ala Ala Gln Pro Phe Leu Leu Arg Asn Gly Ala Asn Glu Gly Phe  
 385                      390                      395                      400  
 His Glu Ala Val Gly Glu Ile Met Ser Leu Ser Ala Ala Thr Pro Lys  
                                  405                      410                      415  
 His Leu Lys Ser Ile Gly Leu Leu Ser Pro Asp Phe Gln Glu Asp Asn  
                                  420                      425                      430  
 Glu Thr Glu Ile Asn Phe Leu Leu Lys Gln Ala Leu Thr Ile Val Gly  
                                  435                      440                      445  
 Thr Leu Pro Phe Thr Tyr Met Leu Glu Lys Trp Arg Trp Met Val Phe  
                                  450                      455                      460  
 Lys Gly Glu Ile Pro Lys Asp Gln Trp Met Lys Lys Trp Trp Glu Met  
 465                      470                      475                      480  
  
 Lys Arg Glu Ile Val Gly Val Val Glu Pro Val Pro His Asp Glu Thr  
                                  485                      490                      495  
 Tyr Cys Asp Pro Ala Ser Leu Phe His Val Ser Asn Asp Tyr Ser Phe

500 505 510  
 Ile Arg Tyr Tyr Thr Arg Thr Leu Tyr Gln Phe Gln Phe Gln Glu Ala  
 515 520 525  
 Leu Cys Gln Ala Ala Lys His Glu Gly Pro Leu His Lys Cys Asp Ile  
 530 535 540  
 Ser Asn Ser Thr Glu Ala Gly Gln Lys Leu Phe Asn Met Leu Arg Leu  
 545 550 555 560  
 Gly Lys Ser Glu Pro Trp Thr Leu Ala Leu Glu Asn Val Val Gly Ala  
 565 570 575  
 Lys Asn Met Asn Val Arg Pro Leu Leu Asn Tyr Phe Glu Pro Leu Phe  
 580 585 590  
 Thr Trp Leu Lys Asp Gln Asn Lys Asn Ser Phe Val Gly Trp Ser Thr  
 595 600 605  
 Asp Trp Ser Pro Tyr Ala Asp Gln Ser Ile Lys Val Arg Ile Ser Leu  
 610 615 620  
 Lys Ser Ala Leu Glu Asp Lys Ala Tyr Glu Trp Asn Asp Asn Glu Met  
 625 630 635 640  
 Tyr Leu Phe Arg Ser Ser Val Ala Tyr Ala Met Arg Gln Tyr Phe Leu  
 645 650 655  
 Lys Val Lys Asn Gln Met Ile Leu Phe Gly Glu Glu Asp Val Arg Val  
 660 665 670  
 Ala Asn Leu Lys Pro Arg Ile Ser Phe Asn Phe Phe Val Thr Ala Pro  
 675 680 685  
 Lys Asn Val Ser Asp Ile Ile Pro Arg Thr Glu Val Glu Lys Ala Ile  
 690 695 700  
 Arg Met Ser Arg Ser Arg Ile Asn Asp Ala Phe Arg Leu Asn Asp Asn  
 705 710 715 720  
 Ser Leu Glu Phe Leu Gly Ile Gln Pro Thr Leu Gly Pro Pro Asn Gln  
 725 730 735  
 Pro Pro Val Ser Ile Trp Leu Ile Val Phe Gly Val Val Met Gly Val  
 740 745 750  
 Ile Val Val Gly Ile Val Ile Leu Ile Phe Thr Gly Ile Arg Asp Arg  
 755 760 765  
 Lys Lys Lys Asn Lys Ala Arg Ser Gly Glu Asn Pro Tyr Ala Ser Ile  
 770 775 780  
 Asp Ile Ser Lys Gly Glu Asn Asn Pro Gly Phe Gln Asn Thr Asp Asp  
 785 790 795 800  
 Val Gln Thr Ser Phe  
 805

&lt;210&gt; 3

&lt;211&gt; 2415

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> This degenerate sequence encodes the amino acid  
sequence of SEQ ID NO:2.

&lt;221&gt; misc\_feature

&lt;222&gt; (1)...(2415)

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 3

atgwsnwsnw	snwsntggyt	nytnytnwsn	ytngtngcng	tnacngcngc	ncarwsnacn	60
athgargarc	argcnaarac	nttytngay	aarttyaayc	aygargcnga	rgayytntty	120
taycarwsnw	snytngcnws	ntggaaytay	aayacnaaya	thacngarga	raaygtncar	180
aayatgaaya	aygcnggnga	yaartggwsn	gcnttyyttna	argarcarws	nacnytngcn	240
caratgtayc	cnytncarga	rathcaraay	ytncngttna	arytncaryt	ncargcnytn	300
carcaraayg	gnwsnwsngt	nytnwsngar	gayaarwsna	armgnytnaa	yacnathytn	360
aayacnatgw	snacnathta	ywsnacnggn	aargtntgya	ayccngayaa	yccncargar	420
tgyytntytn	tngarccngg	nytnaaygar	athatggcna	aywsnytnga	ytayaaygar	480
mgnytnntggg	cntgggarws	ntggmgnwsn	gargtnggna	arcarytnmg	nccnytnntay	540
gargartayg	tngtnytnaa	raaygaratg	gcnmngcna	aycaytayga	rgaytayggg	600
gaytaytggm	gngngayta	ygargtnaay	ggngtngayg	gntaygayta	ywsnmngngn	660
carytnathg	argaygtnga	rcayacntty	gargaratha	arccnytnnta	ygarcayytn	720
caygcntayg	tnmgngcnaa	rytnatgaay	gcntayccnw	sntayathws	nccnathggg	780
tgyytncng	cncayytnty	ngnggayatg	tggggnmgnt	tytggaacnaa	yytnntaywsn	840
ytncngtnc	cnttyggnga	raarccnaay	athgaygttna	cngaygcnat	ggtngaycar	900
gcntgggag	cncarmgnat	httyaargar	gcngaraart	tyttygtngws	ngtnggnytn	960
ccnaayatga	cncarggntt	ytgggaraay	wsnatgytna	cngayccngg	naaygtncar	1020
aargcngtnt	gyayccnac	ngcntgggag	ytnggnaarg	gngayttnmg	nathytnatg	1080
tgyacnaarg	tnacnatgga	ygayttyytn	acngcncayc	aygaratggg	ncayathcar	1140
taygayatgg	cntaygcngc	ncarccntty	ytntymgna	ayggngcnaa	ygarggntty	1200
caygargcng	tngngarat	hatgwsnytn	wsngcngcna	cncncaarca	yytnaarwsn	1260
athggnytny	tnwsnccnga	ytycargar	gayaaygara	cngarathaa	ytyytnytn	1320
aarcargcny	tnacnathgt	nggnacnytn	ccnttyacnt	ayatgytna	raartggmgn	1380
tggatggtnt	tyaarggnga	rathccnaar	gaycartgga	tgaaraartg	gtgggaratg	1440
aarmnggara	thgtngngt	ngtngarccn	gtncncayg	aygaracnta	ytgygayccn	1500
gcnwsnytn	tycaygtngws	naaygaytay	wsnttyathm	gntaytayac	nmgnacnytn	1560
taycarttyc	arttycarga	rgcnytnngy	cargcngcna	arcaygargg	nccnytncaay	1620
aartgygaya	thwsnaayws	nacngargcn	ggncaraary	tnttyaayat	gytnmgnytn	1680
ggnaarwsng	arccntggac	nytnngcnytn	garaaygtng	tngngcnaa	raayatgaay	1740
gtnmngccny	tnytnaayta	ytygarccn	ytnttyacnt	ggytnaarga	ycaraayaar	1800
aaywsnttyg	tnggntgqws	nacngaytgg	wsnccntayg	cngaycarws	nathaargtn	1860
mgnathwsny	tnaarwsngc	nytnngngay	aargcntayg	artggaayga	yaaygaratg	1920
tayytntty	gnwsnwsngt	ngcntaygc	atgmngcart	ayttyytnaa	rgtnaaraay	1980
caratgathy	tnttyggnga	rgargaygt	mgngtngcna	ayytnaarcc	nmgnathwsn	2040
ttyaaytyt	tygtnacngc	nccnaaraay	gtngwsngaya	thathccnmg	nacngargtn	2100
garaargcna	thmgntatgws	nmgnwsnmgn	athaaygayg	cnttymgny	naaygayaay	2160

wsnytngart tyytnggnat hcarccnacn ytnggncnc cnaaycarcc nccngtnwsn	2220
athtggytna thgtnttygg ngtngtnatg ggngtnathg tngtnggnat hgtgnathyt	2280
athttyacng gnathmgnga ymgnaaraar aaraayaarg cnmgwnsgg ngaraayccn	2340
taygcwnsna thgayathws naarggngar aayaayccng gnttycaraa yacngaygay	2400
gtncaracnw snnty	2415

<210> 4  
 <211> 16  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Peptide linker

<400> 4  
 Gly Gly Ser Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser  
 1 5 10 15

<210> 5  
 <211> 2638  
 <212> DNA  
 <213> Mouse

<220>  
 <221> CDS  
 <222> (106)...(2520)

<400> 5  
 agtgcccaac ccaagttcaa aggctgatga gagagaaaaa ctcatgaaga gattttactc 60  
 tagggaaagt tgctcagtgg atgggatctt ggcgcacggg gaaag atg tcc agc tcc 117  
 Met Ser Ser Ser  
 1

tcc tgg ctc ctt ctc agc ctt gtt gct gtt act act gct cag tcc ctc 165  
 Ser Trp Leu Leu Leu Ser Leu Val Ala Val Thr Thr Ala Gln Ser Leu  
 5 10 15 20

acc gag gaa aat gcc aag aca ttt tta aac aac ttt aat cag gaa gct 213  
 Thr Glu Glu Asn Ala Lys Thr Phe Leu Asn Asn Phe Asn Gln Glu Ala  
 25 30 35

gaa gac ctg tct tat caa agt tca ctt gct tct tgg aat tat aat act 261  
 Glu Asp Leu Ser Tyr Gln Ser Ser Leu Ala Ser Trp Asn Tyr Asn Thr  
 40 45 50



aac att act gaa gaa aat gcc caa aag atg agt gag gct gca gcc aaa 309  
 Asn Ile Thr Glu Glu Asn Ala Gln Lys Met Ser Glu Ala Ala Ala Lys  
 55 60 65

tgg tct gcc ttt tat gaa gaa cag tct aag act gcc caa agt ttc tca 357  
 Trp Ser Ala Phe Tyr Glu Glu Gln Ser Lys Thr Ala Gln Ser Phe Ser  
 70 75 80

cta caa gaa atc cag act ccg atc atc aag cgt caa cta cag gcc ctt 405  
 Leu Gln Glu Ile Gln Thr Pro Ile Ile Lys Arg Gln Leu Gln Ala Leu  
 85 90 95 100

cag caa agt ggg tct tca gca ctc tca gca gac aag aac aaa cag ttg 453  
 Gln Gln Ser Gly Ser Ser Ala Leu Ser Ala Asp Lys Asn Lys Gln Leu  
 105 110 115

aac aca att ctg aac acc atg agc acc att tac agt act gga aaa gtt 501  
 Asn Thr Ile Leu Asn Thr Met Ser Thr Ile Tyr Ser Thr Gly Lys Val  
 120 125 130

tgc aac cca aag aac cca caa gaa tgc tta tta ctt gag cca gga ttg 549  
 Cys Asn Pro Lys Asn Pro Gln Glu Cys Leu Leu Leu Glu Pro Gly Leu  
 135 140 145

gat gaa ata atg gcg aca agc aca gac tac aac tct agg ctc tgg gca 597  
 Asp Glu Ile Met Ala Thr Ser Thr Asp Tyr Asn Ser Arg Leu Trp Ala  
 150 155 160

tgg gag ggc tgg agg gct gag gtt ggc aag cag ctg agg ccg ttg tat 645  
 Trp Glu Gly Trp Arg Ala Glu Val Gly Lys Gln Leu Arg Pro Leu Tyr  
 165 170 175 180

gaa gag tat gtg gtc ctg aaa aac gag atg gca aga gca aac aat tat 693  
 Glu Glu Tyr Val Val Leu Lys Asn Glu Met Ala Arg Ala Asn Asn Tyr  
 185 190 195

aac gac tat ggg gat tat tgg aga ggg gac tat gaa gca gag gga gca 741  
 Asn Asp Tyr Gly Asp Tyr Trp Arg Gly Asp Tyr Glu Ala Glu Gly Ala  
 200 205 210

gat ggc tac aac tat aac cgt aac cag ttg att gaa gat gta gaa cgt 789  
 Asp Gly Tyr Asn Tyr Asn Arg Asn Gln Leu Ile Glu Asp Val Glu Arg  
 215 220 225

acc ttc gca gag atr aag cca ttg tat gag cat ctt cat gcc tat gtg Thr Phe Ala Glu Ile Lys Pro Leu Tyr Glu His Leu His Ala Tyr Val 230 235 240	837
agg agg aag ttg atg gat acc tac cct tcc tac atc agc ccc act gga Arg Arg Lys Leu Met Asp Thr Tyr Pro Ser Tyr Ile Ser Pro Thr Gly 245 250 255 260	885
tgc ctc cct gcc cat ttg ctt ggt gat atg tgg ggt aga ttt tgg aca Cys Leu Pro Ala His Leu Leu Gly Asp Met Trp Gly Arg Phe Trp Thr 265 270 275	933
aat ctg tac cct ttg act gtt ccc ttt gca cag aaa cca aac ata gat Asn Leu Tyr Pro Leu Thr Val Pro Phe Ala Gln Lys Pro Asn Ile Asp 280 285 290	981
gtt act gat gca atg atg aat cag ggc tgg gat gca gaa agg ata ttt Val Thr Asp Ala Met Met Asn Gln Gly Trp Asp Ala Glu Arg Ile Phe 295 300 305	1029
caa gag gca gag aaa ttc ttt gtt tct gtt ggc ctt cct cat atg act Gln Glu Ala Glu Lys Phe Phe Val Ser Val Gly Leu Pro His Met Thr 310 315 320	1077
caa gga ttc tgg gca aac tct atg ctg act gag cca gca gat ggc cgg Gln Gly Phe Trp Ala Asn Ser Met Leu Thr Glu Pro Ala Asp Gly Arg 325 330 335 340	1125
aaa gtt gtc tgc cac ccc aca gct tgg gat ctg gga cac gga gac ttc Lys Val Val Cys His Pro Thr Ala Trp Asp Leu Gly His Gly Asp Phe 345 350 355	1173
aga atc aag atg tgt aca aag gtc aca atg gac aac ttc ttg aca gcc Arg Ile Lys Met Cys Thr Lys Val Thr Met Asp Asn Phe Leu Thr Ala 360 365 370	1221
cat cac gag atg gga cac atc caa tat gac atg gca tat gcc agg caa His His Glu Met Gly His Ile Gln Tyr Asp Met Ala Tyr Ala Arg Gln 375 380 385	1269
cct ttc ctg cta aga aac gga gcc aat gaa ggg ttc cat gaa gct gtt Pro Phe Leu Leu Arg Asn Gly Ala Asn Glu Gly Phe His Glu Ala Val 390 395 400	1317

gga gaa atc atg tca ctt tct gca gct acc ccc aag cat ctg aaa tcc Gly Glu Ile Met Ser Leu Ser Ala Ala Thr Pro Lys His Leu Lys Ser 405 410 415 420	1365
att ggt ctt ctg cca tcc gat ttt caa gaa gat agc gaa aca gag ata Ile Gly Leu Leu Pro Ser Asp Phe Gln Glu Asp Ser Glu Thr Glu Ile 425 430 435	1413
aac ttc cta ctg aaa cag gca ttg aca att gtt gga aca cta ccg ttt Asn Phe Leu Leu Lys Gln Ala Leu Thr Ile Val Gly Thr Leu Pro Phe 440 445 450	1461
act tac atg tta gag aag tgg agg tgg atg gtc ttt cgg ggt gaa att Thr Tyr Met Leu Glu Lys Trp Arg Trp Met Val Phe Arg Gly Glu Ile 455 460 465	1509
ccc aaa gag cag tgg atg aaa aag tgg tgg gag atg aag cgg gag atc Pro Lys Glu Gln Trp Met Lys Lys Trp Trp Glu Met Lys Arg Glu Ile 470 475 480	1557
gtt ggt gtg gtg gag cct ctg cct cat gat gaa aca tac tgt gac cct Val Gly Val Val Glu Pro Leu Pro His Asp Glu Thr Tyr Cys Asp Pro 485 490 495 500	1605
gca tct ctg ttc cat gtt tct aat gat tac tca ttc att cga tat tac Ala Ser Leu Phe His Val Ser Asn Asp Tyr Ser Phe Ile Arg Tyr Tyr 505 510 515	1653
aca agg acc att tac caa ttc cag ttt caa gaa gct ctt tgt caa gca Thr Arg Thr Ile Tyr Gln Phe Gln Phe Gln Glu Ala Leu Cys Gln Ala 520 525 530	1701
gct aag tat aat ggt tct ctg cac aaa tgt gac atc tca aat tcc act Ala Lys Tyr Asn Gly Ser Leu His Lys Cys Asp Ile Ser Asn Ser Thr 535 540 545	1749
gaa gct ggg cag aag ttg ctc aag atg ctg agt ctt gga aat tca gag Glu Ala Gly Gln Lys Leu Leu Lys Met Leu Ser Leu Gly Asn Ser Glu 550 555 560	1797
ccc tgg acc aaa gcc ttg gaa aat gtg gta gga gca agg aat atg gat Pro Trp Thr Lys Ala Leu Glu Asn Val Val Gly Ala Arg Asn Met Asp 565 570 575 580	1845

gta aaa cca ctg ctc aat tac ttc caa ccg ttg ttt gac tgg ctg aaa Val Lys Pro Leu Leu Asn Tyr Phe Gln Pro Leu Phe Asp Trp Leu Lys 585 590 595	1893
gag cag aac aga aat tct ttt gtg ggg tgg aac act gaa tgg agc cca Glu Gln Asn Arg Asn Ser Phe Val Gly Trp Asn Thr Glu Trp Ser Pro 600 605 610	1941
tat gcc gac caa agc att aaa gtg agg ata agc cta aaa tca gct ctt Tyr Ala Asp Gln Ser Ile Lys Val Arg Ile Ser Leu Lys Ser Ala Leu 615 620 625	1989
gga gct aat gca tat gaa tgg acc aac aac gaa atg ttc ctg ttc cga Gly Ala Asn Ala Tyr Glu Trp Thr Asn Asn Glu Met Phe Leu Phe Arg 630 635 640	2037
tca tct gtt gca tat gcc atg aga aag tat ttt tca ata atc aaa aac Ser Ser Val Ala Tyr Ala Met Arg Lys Tyr Phe Ser Ile Ile Lys Asn 645 650 655 660	2085
cag aca gtt cct ttt cta gag gaa gat gta cga gtg agc gat ttg aaa Gln Thr Val Pro Phe Leu Glu Glu Asp Val Arg Val Ser Asp Leu Lys 665 670 675	2133
cca aga gtc tcc ttc tac ttc ttt gtc acc tca ccc caa aat gtg tct Pro Arg Val Ser Phe Tyr Phe Phe Val Thr Ser Pro Gln Asn Val Ser 680 685 690	2181
gat gtc att cct aga agt gaa gtt gaa gat gcc atc agg atg tct cgg Asp Val Ile Pro Arg Ser Glu Val Glu Asp Ala Ile Arg Met Ser Arg 695 700 705	2229
ggc cgc atc aat gat gtc ttt ggc ctg aat gat aac agc ctg gag ttt Gly Arg Ile Asn Asp Val Phe Gly Leu Asn Asp Asn Ser Leu Glu Phe 710 715 720	2277
ctg ggg att cac cca aca ctt gag cca cct tac cag cct cct gtc acc Leu Gly Ile His Pro Thr Leu Glu Pro Pro Tyr Gln Pro Pro Val Thr 725 730 735 740	2325
ata tgg ctg att att ttt ggt gtt gtg atg gca ctg gta gtg gtt ggc Ile Trp Leu Ile Ile Phe Gly Val Val Met Ala Leu Val Val Val Gly 745 750 755	2373

atc atc atc ctg att gtc act ggg atc aaa ggt cga aag aag aaa aat 2421  
 Ile Ile Ile Leu Ile Val Thr Gly Ile Lys Gly Arg Lys Lys Lys Asn  
 760 765 770

gaa aca aaa aga gaa gag aac cct tat gac tcg atg gac att gga aaa 2469  
 Glu Thr Lys Arg Glu Glu Asn Pro Tyr Asp Ser Met Asp Ile Gly Lys  
 775 780 785

gga gaa agc aat gca gga ttc caa aac agt gat gat gct cag act tcc 2517  
 Gly Glu Ser Asn Ala Gly Phe Gln Asn Ser Asp Asp Ala Gln Thr Ser  
 790 795 800

ttt tagcaaagca cttgtcatct tcctgtatgt aaatgctaac ttcatagtac 2570  
 Phe  
 805

acaaaatatg agagtataca catgtcatta gctatcaaaa ctatgatctg ttcagtaaac 2630  
 gttgtcca 2638

<210> 6

<211> 805

<212> PRT

<213> Mouse

<400> 6

Met Ser Ser Ser Ser Trp Leu Leu Leu Ser Leu Val Ala Val Thr Thr  
 1 5 10 15  
 Ala Gln Ser Leu Thr Glu Glu Asn Ala Lys Thr Phe Leu Asn Asn Phe  
 20 25 30  
 Asn Gln Glu Ala Glu Asp Leu Ser Tyr Gln Ser Ser Leu Ala Ser Trp  
 35 40 45  
 Asn Tyr Asn Thr Asn Ile Thr Glu Glu Asn Ala Gln Lys Met Ser Glu  
 50 55 60  
 Ala Ala Ala Lys Trp Ser Ala Phe Tyr Glu Glu Gln Ser Lys Thr Ala  
 65 70 75 80  
 Gln Ser Phe Ser Leu Gln Glu Ile Gln Thr Pro Ile Ile Lys Arg Gln  
 85 90 95  
 Leu Gln Ala Leu Gln Gln Ser Gly Ser Ser Ala Leu Ser Ala Asp Lys  
 100 105 110  
 Asn Lys Gln Leu Asn Thr Ile Leu Asn Thr Met Ser Thr Ile Tyr Ser  
 115 120 125

Thr Gly Lys Val Cys Asn Pro Lys Asn Pro Gln Glu Cys Leu Leu Leu  
 130 135 140  
 Glu Pro Gly Leu Asp Glu Ile Met Ala Thr Ser Thr Asp Tyr Asn Ser

145                      150                      155                      160  
 Arg Leu Trp Ala Trp Glu Gly Trp Arg Ala Glu Val Gly Lys Gln Leu  
                                  165                      170                      175  
 Arg Pro Leu Tyr Glu Glu Tyr Val Val Leu Lys Asn Glu Met Ala Arg  
                                  180                      185                      190  
 Ala Asn Asn Tyr Asn Asp Tyr Gly Asp Tyr Trp Arg Gly Asp Tyr Glu  
                                  195                      200                      205  
 Ala Glu Gly Ala Asp Gly Tyr Asn Tyr Asn Arg Asn Gln Leu Ile Glu  
                                  210                      215                      220  
 Asp Val Glu Arg Thr Phe Ala Glu Ile Lys Pro Leu Tyr Glu His Leu  
 225                      230                      235                      240  
 His Ala Tyr Val Arg Arg Lys Leu Met Asp Thr Tyr Pro Ser Tyr Ile  
                                  245                      250                      255  
 Ser Pro Thr Gly Cys Leu Pro Ala His Leu Leu Gly Asp Met Trp Gly  
                                  260                      265                      270  
 Arg Phe Trp Thr Asn Leu Tyr Pro Leu Thr Val Pro Phe Ala Gln Lys  
                                  275                      280                      285  
 Pro Asn Ile Asp Val Thr Asp Ala Met Met Asn Gln Gly Trp Asp Ala  
                                  290                      295                      300  
 Glu Arg Ile Phe Gln Glu Ala Glu Lys Phe Phe Val Ser Val Gly Leu  
 305                      310                      315                      320  
 Pro His Met Thr Gln Gly Phe Trp Ala Asn Ser Met Leu Thr Glu Pro  
                                  325                      330                      335  
 Ala Asp Gly Arg Lys Val Val Cys His Pro Thr Ala Trp Asp Leu Gly  
                                  340                      345                      350  
 His Gly Asp Phe Arg Ile Lys Met Cys Thr Lys Val Thr Met Asp Asn  
                                  355                      360                      365  
 Phe Leu Thr Ala His His Glu Met Gly His Ile Gln Tyr Asp Met Ala  
                                  370                      375                      380  
 Tyr Ala Arg Gln Pro Phe Leu Leu Arg Asn Gly Ala Asn Glu Gly Phe  
 385                      390                      395                      400  
 His Glu Ala Val Gly Glu Ile Met Ser Leu Ser Ala Ala Thr Pro Lys  
                                  405                      410                      415  
 His Leu Lys Ser Ile Gly Leu Leu Pro Ser Asp Phe Gln Glu Asp Ser  
                                  420                      425                      430  
 Glu Thr Glu Ile Asn Phe Leu Leu Lys Gln Ala Leu Thr Ile Val Gly  
                                  435                      440                      445  
 Thr Leu Pro Phe Thr Tyr Met Leu Glu Lys Trp Arg Trp Met Val Phe  
                                  450                      455                      460  
 Arg Gly Glu Ile Pro Lys Glu Gln Trp Met Lys Lys Trp Trp Glu Met  
 465                      470                      475                      480  
  
 Lys Arg Glu Ile Val Gly Val Val Glu Pro Leu Pro His Asp Glu Thr  
                                  485                      490                      495  
 Tyr Cys Asp Pro Ala Ser Leu Phe His Val Ser Asn Asp Tyr Ser Phe

500 505 510  
 Ile Arg Tyr Tyr Thr Arg Thr Ile Tyr Gln Phe Gln Phe Gln Glu Ala  
 515 520 525  
 Leu Cys Gln Ala Ala Lys Tyr Asn Gly Ser Leu His Lys Cys Asp Ile  
 530 535 540  
 Ser Asn Ser Thr Glu Ala Gly Gln Lys Leu Leu Lys Met Leu Ser Leu  
 545 550 555 560  
 Gly Asn Ser Glu Pro Trp Thr Lys Ala Leu Glu Asn Val Val Gly Ala  
 565 570 575  
 Arg Asn Met Asp Val Lys Pro Leu Leu Asn Tyr Phe Gln Pro Leu Phe  
 580 585 590  
 Asp Trp Leu Lys Glu Gln Asn Arg Asn Ser Phe Val Gly Trp Asn Thr  
 595 600 605  
 Glu Trp Ser Pro Tyr Ala Asp Gln Ser Ile Lys Val Arg Ile Ser Leu  
 610 615 620  
 Lys Ser Ala Leu Gly Ala Asn Ala Tyr Glu Trp Thr Asn Asn Glu Met  
 625 630 635 640  
 Phe Leu Phe Arg Ser Ser Val Ala Tyr Ala Met Arg Lys Tyr Phe Ser  
 645 650 655  
 Ile Ile Lys Asn Gln Thr Val Pro Phe Leu Glu Glu Asp Val Arg Val  
 660 665 670  
 Ser Asp Leu Lys Pro Arg Val Ser Phe Tyr Phe Phe Val Thr Ser Pro  
 675 680 685  
 Gln Asn Val Ser Asp Val Ile Pro Arg Ser Glu Val Glu Asp Ala Ile  
 690 695 700  
 Arg Met Ser Arg Gly Arg Ile Asn Asp Val Phe Gly Leu Asn Asp Asn  
 705 710 715 720  
 Ser Leu Glu Phe Leu Gly Ile His Pro Thr Leu Glu Pro Pro Tyr Gln  
 725 730 735  
 Pro Pro Val Thr Ile Trp Leu Ile Ile Phe Gly Val Val Met Ala Leu  
 740 745 750  
 Val Val Val Gly Ile Ile Ile Leu Ile Val Thr Gly Ile Lys Gly Arg  
 755 760 765  
 Lys Lys Lys Asn Glu Thr Lys Arg Glu Glu Asn Pro Tyr Asp Ser Met  
 770 775 780  
 Asp Ile Gly Lys Gly Glu Ser Asn Ala Gly Phe Gln Asn Ser Asp Asp  
 785 790 795 800  
 Ala Gln Thr Ser Phe  
 805

&lt;210&gt; 7

&lt;211&gt; 2415

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> This degenerate sequence encodes the amino acid  
sequence of SEQ ID NO:6.

&lt;221&gt; misc\_feature

&lt;222&gt; (1)...(2415)

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 7

atgwsnwsnw	snwsntggyt	nytnytnwsn	ytngtngcng	tnacnacngc	ncarwsnytn	60
acngargara	aygcnaarac	nttyytnaay	aayttyaayc	argargcnga	rgayytwnsn	120
taycarwsnw	snytngcnws	ntggaaytay	aayacnaaya	thacngarga	raaygcncar	180
aratgwsng	argcngcngc	naartggwsn	gcnttytayg	argarcarws	naaracngcn	240
carwsnttyw	snytncarga	rathcaracn	ccnathatha	armgncaryt	ncargcnytn	300
carcarwsng	gnwsnwsngc	nytnwsngcn	gayaaraaya	arcarytnaa	yacnathytn	360
aayacnatgw	snacnatnta	ywsnacnggn	aargtntgya	ayccnaaraa	yccncargar	420
tgyytnytny	tngarccngg	nytngaygar	athatggcna	cnwsnacnga	ytayaaywsn	480
mgnytnthgg	cntgggargg	ntggmgngcn	gargtnggna	arcarytnmg	nccnytnay	540
gargartayg	tngtnytnaa	raaygaratg	gcnmngngcna	ayaaytayaa	ygaytayggg	600
gaytaytggm	gnngngayta	ygargcngar	ggngcngayg	gntayaayta	yaaymgnaay	660
carytnathg	argaytgnga	rmgnacntty	gcngaratha	arccnytna	ygarcayytn	720
caaycmtayg	tnmgngmnaa	rytnatggay	acntayccnw	sntayathws	nccnacnggn	780
tgyytnccng	cncayytnyt	ngngngaytg	tggggngmgt	tytggaacnaa	yytnayccn	840
ytncangtnc	cnttygcnc	raarccnaay	athgaytgna	cngaygcnat	gatgaaycar	900
ggntgggag	cngarmgnat	httycargar	gcngaraart	tyttygtngs	ngtnggnytn	960
ccncayatga	cncarggntt	ytgggcnaay	wsnatgytna	cngarccngc	ngayggngmn	1020
aargtngtnt	gyayccnac	ngcntgggag	ytnggncayg	gngayttymg	nathaaratg	1080
tgyacnaarg	tnacnatgga	yaayttyytn	acngcncayc	aygaratggg	ncayathcar	1140
taygayatgg	cntaygcmmg	ncarccntty	ytntymgna	ayggngcnaa	ygarggntty	1200
caygargcng	tngnggarat	hatgwsnytn	wsngcngcna	cncnaarca	yytnaarwsn	1260
athggnytny	tnccnwsnga	yttycargar	gaywsngara	cngarathaa	yttyytnytn	1320
aarcargcny	tnacnathgt	nggnacnytn	ccnttyacnt	ayatgytnga	raartggmgn	1380
tggatggtnt	tymngngnga	rathccnaar	garcartgga	tgaaraartg	gtgggaratg	1440
aarmnggara	thgtngnggt	ngtngarccn	ytncncayg	aygaracnta	ytgygayccn	1500
gcnwsnytn	tycaygtngs	naaygaytay	wsnttyathm	gntaytayac	nmgnacnath	1560
taycarttyc	arttycarga	rgcnyntgy	cargcngcna	artayaaygg	nwsnytnca	1620
aartgygaya	thwsnaayws	nacngargcn	ggncaraary	tnytnaarat	gytnwsnytn	1680
ggnaaywsng	arccntggac	naargcnytn	garaaygtng	tnggngcnmg	naayatggay	1740
gtnaarccny	tnytnaayta	yttycarccn	ytnttygayt	ggytnaarga	rcaraaymgn	1800
aaywsnttyg	tnggntggaa	yacngartgg	wsnccntayg	cngaycarws	nathargtn	1860
mgnathwsny	tnaarwsngc	nytnngngcn	aaygcmtayg	artggacnaa	yaaygaratg	1920
ttyytnntym	gnwsnwsngt	ngcntaygc	atgmngnaart	ayttywsnat	hathaaraay	1980
caracngtnc	cnttyytnga	rgargaygt	mgngtnwsng	ayytnaarcc	nmngtnwsn	2040
ttytayttyt	tygtnacnws	nccncaraay	gtngwsngayg	tnathccnmg	nwsngargtn	2100
gargaygcna	thmgngatgws	nmngngnmgn	athaaygayg	tnnttyggny	naaygayaay	2160



wsnytngart tyytnggnat hcayccnacn ytngarccnc cntaycarcc nccngtnacn	2220
athtggytna thathtvvg ngtngtnatg gcnytngtng tngtnggnat hathathytn	2280
athgtnacng gnathaaagg nmgnaraar aaraaygara cnaarmgnga rgaraayccn	2340
taggaywsna tggayathgg naargngar wsnaaygcng gnttycaraa ywsngaygay	2400
gcncaracnw snTTY	2415

&lt;210&gt; 8

&lt;211&gt; 2638

&lt;212&gt; DNA

&lt;213&gt; Mouse

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (106)...(2520)

&lt;400&gt; 8

agtcccaac ccaagttcaa aggctgatga gagagaaaaa ctcatgaaga gattttactc	60
tagggaaagt tgctcagtgg atgggatctt ggcgacggg gaaag atg tcc agc tcc	117
Met Ser Ser Ser	
1	

tcc tgg ctc ctt ctc agc ctt gtt gct gtt act act gct cag tcc ctc	165
Ser Trp Leu Leu Leu Ser Leu Val Ala Val Thr Thr Ala Gln Ser Leu	
5 10 15 20	

acc gag gaa aat gcc aag aca ttt tta aac aac ttt aat cag gag gct	213
Thr Glu Glu Asn Ala Lys Thr Phe Leu Asn Asn Phe Asn Gln Glu Ala	
25 30 35	

gaa gac ctg tct tat caa agt tca ctt gct tct tgg aat tat aat act	261
Glu Asp Leu Ser Tyr Gln Ser Ser Leu Ala Ser Trp Asn Tyr Asn Thr	
40 45 50	

aac att act gaa gaa aat gcc caa aag atg agt gag gct gca gcc aaa	309
Asn Ile Thr Glu Glu Asn Ala Gln Lys Met Ser Glu Ala Ala Ala Lys	
55 60 65	

tgg tct gcc ttt tat gaa gaa cag tct aag act gcc caa agt ttc tca	357
Trp Ser Ala Phe Tyr Glu Glu Gln Ser Lys Thr Ala Gln Ser Phe Ser	
70 75 80	

cta caa gaa atc cag act ccg atc atc aag cgt caa cta cag gcc ctt	405
Leu Gln Glu Ile Gln Thr Pro Ile Ile Lys Arg Gln Leu Gln Ala Leu	
85 90 95 100	

cag caa agt ggg tct tca gca ctc tca gca gac aag aac aaa cag ttg 453  
 Gln Gln Ser Gly Ser Ser Ala Leu Ser Ala Asp Lys Asn Lys Gln Leu  
 105 110 115

aac aca att ctg aac acc atg agc acc att tac agt act gga aaa gtt 501  
 Asn Thr Ile Leu Asn Thr Met Ser Thr Ile Tyr Ser Thr Gly Lys Val  
 120 125 130

tgc aac cca agg aac cca caa gaa tgc tta tta ctt gag cca gga ttg 549  
 Cys Asn Pro Arg Asn Pro Gln Glu Cys Leu Leu Leu Glu Pro Gly Leu  
 135 140 145

gat gaa ata atg gcg aca agc aca gac tac aac tct agg ctc tgg gca 597  
 Asp Glu Ile Met Ala Thr Ser Thr Asp Tyr Asn Ser Arg Leu Trp Ala  
 150 155 160

tgg gag ggc tgg agg gct gag gtt ggc aag cag ctg agg ccg ttg tat 645  
 Trp Glu Gly Trp Arg Ala Glu Val Gly Lys Gln Leu Arg Pro Leu Tyr  
 165 170 175 180

gaa gag tat gtg gtc ctg aaa aac gag atg gca aga gca aac aat tat 693  
 Glu Glu Tyr Val Val Leu Lys Asn Glu Met Ala Arg Ala Asn Asn Tyr  
 185 190 195

aac gac tat ggg gat tat tgg aga ggg gac tat gaa gca gag gga gca 741  
 Asn Asp Tyr Gly Asn Tyr Trp Arg Gly Asp Tyr Glu Ala Glu Gly Ala  
 200 205 210

gat ggc tac aac tat aac cgt aac cag ttg att gaa gat gta gaa cgt 789  
 Asp Gly Tyr Asn Tyr Asn Arg Asn Gln Leu Ile Glu Asp Val Glu Arg  
 215 220 225

acc ttc gca gag atc aag cca ttg tat gag cat ctt cat gcc tat gtg 837  
 Thr Phe Ala Glu Ile Lys Pro Leu Tyr Glu His Leu His Ala Tyr Val  
 230 235 240

agg agg aag ttg atg gat acc tac cct tcc tac atc agc ccc act gga 885  
 Arg Arg Lys Leu Met Asp Thr Tyr Pro Ser Tyr Ile Ser Pro Thr Gly  
 245 250 255 260

tgc ctc cct gcc cat ttg ctt ggt gat atg tgg ggt aga ttt tgg aca 933  
 Cys Leu Pro Ala His Leu Leu Gly Asp Met Trp Gly Arg Phe Trp Thr  
 265 270 275

aat ctg tac cct ttg act gtt ccc ttt gca cag aaa cca aac ata gat 981  
 Asn Leu Tyr Pro Leu Thr Val Pro Phe Ala Gln Lys Pro Asn Ile Asp  
                   280                                  285                                  290

gtt act gat gca atg atg aat cag ggc tgg gat gca gaa agg ata ttt 1029  
 Val Thr Asp Ala Met Met Asn Gln Gly Trp Asp Ala Glu Arg Ile Phe  
                   295                                  300                                  305

caa gag gca gag aaa ttc ttt gtt tct gtt ggc ctt cct cat atg act 1077  
 Gln Glu Ala Glu Lys Phe Phe Val Ser Val Gly Leu Pro His Met Thr  
                   310                                  315                                  320

caa gga ttc tgg gca aac tct atg ctg act gag cca gca gat ggc cgg 1125  
 Gln Gly Phe Trp Ala Asn Ser Met Leu Thr Glu Pro Ala Asp Gly Arg  
                   325                                  330                                  335                                  340

aaa gtt gtc tgc cac ccc aca gct tgg gat ctg gga cac gga gac ttc 1173  
 Lys Val Val Cys His Pro Thr Ala Trp Asp Leu Gly His Gly Asp Phe  
                                   345                                  350                                  355

aga atc aag atg tgt aca aag gtc aca atg gac aac ttc ttg aca gcc 1221  
 Arg Ile Lys Met Cys Thr Lys Val Thr Met Asp Asn Phe Leu Thr Ala  
                   360                                  365                                  370

cat cac gag atg gga cac atc caa tat gac atg gca tat gcc agg caa 1269  
 His His Glu Met Gly His Ile Gln Tyr Asp Met Ala Tyr Ala Arg Gln  
                   375                                  380                                  385

cct ttc ctg cta aga aac gga gcc aat gaa ggg ttc cat gaa gct gtt 1317  
 Pro Phe Leu Leu Arg Asn Gly Ala Asn Glu Gly Phe His Glu Ala Val  
                   390                                  395                                  400

gga gaa atc atg tca ctt tct gca gct acc ccc aag cat ctg aaa tcc 1365  
 Gly Glu Ile Met Ser Leu Ser Ala Ala Thr Pro Lys His Leu Lys Ser  
                   405                                  410                                  415                                  420

att ggt ctt ctg cca tcc gat ttt caa gaa gat agc gaa aca gag ata 1413  
 Ile Gly Leu Leu Pro Ser Asp Phe Gln Glu Asp Ser Glu Thr Glu Ile  
                   425                                  430                                  435

aac ttc cta ctg aaa cag gca ttg aca att gtt gga aca cta ccg ttt 1461  
 Asn Phe Leu Leu Lys Gln Ala Leu Thr Ile Val Gly Thr Leu Pro Phe  
                   440                                  445                                  450

act tac atg tta gag aag tgg agg tgg atg gtc ttt cgg ggt gaa att	1509
Thr Tyr Met Leu Glu Lys Trp Arg Trp Met Val Phe Arg Gly Glu Ile	
455 460 465	
ccc aaa gag cag tgg atg aaa aag tgg tgg gag atg aag cgg gag atc	1557
Pro Lys Glu Gln Trp Met Lys Lys Trp Trp Glu Met Lys Arg Glu Ile	
470 475 480	
gtt ggt gtg gtg gag cct ctg cct cgt gat gaa aca tac tgt gac cct	1605
Val Gly Val Val Glu Pro Leu Pro Arg Asp Glu Thr Tyr Cys Asp Pro	
485 490 495 500	
gca tct ctg ttc cat gtt tct aat gat tac tca ttc att cga tat tac	1653
Ala Ser Leu Phe His Val Ser Asn Asp Tyr Ser Phe Ile Arg Tyr Tyr	
505 510 515	
aca agg acc att tac caa ttc cag ttt caa gaa gct ctt tgt caa gca	1701
Thr Arg Thr Ile Tyr Gln Phe Gln Phe Gln Glu Ala Leu Cys Gln Ala	
520 525 530	
gct aag tat aat ggt tct ctg cac aaa tgt gac atc tca aat tcc act	1749
Ala Lys Tyr Asn Gly Ser Leu His Lys Cys Asp Ile Ser Asn Ser Thr	
535 540 545	
gaa gct ggg cag aag ttg ctc aag atg ctg agt ctt gga aat tca gag	1797
Glu Ala Gly Gln Lys Leu Leu Lys Met Leu Ser Leu Gly Asn Ser Glu	
550 555 560	
ccc tgg acc gaa gcc ttg gaa aat gtg gta gga gca agg aat atg gat	1845
Pro Trp Thr Glu Ala Leu Glu Asn Val Val Gly Ala Arg Asn Met Asp	
565 570 575 580	
gta aaa cca ctg ctc aat tac ttc caa ccg ttg ttt gac tgg ctg aaa	1893
Val Lys Pro Leu Leu Asn Tyr Phe Gln Pro Leu Phe Asp Trp Leu Lys	
585 590 595	
gag cag aac aga aat tct ttt gtg ggg tgg aac act gaa tgg agc cca	1941
Glu Gln Asn Arg Asn Ser Phe Val Gly Trp Asn Thr Glu Trp Ser Pro	
600 605 610	
tat gcc gac caa agc att aaa gtg agg ata agc cta aaa tca gct ctt	1989
Tyr Ala Asp Gln Ser Ile Lys Val Arg Ile Ser Leu Lys Ser Ala Leu	
615 620 625	

gga gct aat gca tat gaa tgg acc aac aac gaa atg ttc ctg ttc cga Gly Ala Asn Ala Tyr Glu Trp Thr Asn Asn Glu Met Phe Leu Phe Arg 630 635 640	2037
tca tct gtt gca tat gcc atg aga aag tat tct tca ata atc aaa aac Ser Ser Val Ala Tyr Ala Met Arg Lys Tyr Ser Ser Ile Ile Lys Asn 645 650 655 660	2085
cag aca gtt cct ttt cta gag gaa gat gta cga gtg agt gat ttg aaa Gln Thr Val Pro Phe Leu Glu Glu Asp Val Arg Val Ser Asp Leu Lys 665 670 675	2133
cca aga gtc tcc ttc tac ttc ttt gtc acc tca ccc caa aat gtg tct Pro Arg Val Ser Phe Tyr Phe Phe Val Thr Ser Pro Gln Asn Val Ser 680 685 690	2181
gat gtc att cct aga agt gaa gtt gaa gat gcc atc agg atg tct cgg Asp Val Ile Pro Arg Ser Glu Val Glu Asp Ala Ile Arg Met Ser Arg 695 700 705	2229
ggc cgc atc aat gat gtc ttt ggc ctg aat gat aac agc ctg gag ttt Gly Arg Ile Asn Asp Val Phe Gly Leu Asn Asp Asn Ser Leu Glu Phe 710 715 720	2277
ctg ggg att cac cca aca ctt gag cca cct tac cag cct cct gtc acc Leu Gly Ile His Pro Thr Leu Glu Pro Pro Tyr Gln Pro Pro Val Thr 725 730 735 740	2325
ata tgg ctg att att ttt ggt gtt gtg atg gca ctg gta gtg gtt ggc Ile Trp Leu Ile Ile Phe Gly Val Val Met Ala Leu Val Val Val Gly 745 750 755	2373
atc atc atc ctg att gtc act ggg atc aaa ggt cga aag aag aaa aat Ile Ile Ile Leu Ile Val Thr Gly Ile Lys Gly Arg Lys Lys Lys Asn 760 765 770	2421
gaa aca aaa aga gaa gag aac cct tat gac tcg atg gac att gga aaa Glu Thr Lys Arg Glu Glu Asn Pro Tyr Asp Ser Met Asp Ile Gly Lys 775 780 785	2469
gga gaa agc aat gca gga ttc caa aac agt gat gat gct cag act tcc Gly Glu Ser Asn Ala Gly Phe Gln Asn Ser Asp Asp Ala Gln Thr Ser 790 795 800	2517

ttt tagcaaagca cttgtcatct tcctgtatgt aaatgctaac ttcatagtac 2570

Phe

805

acaaaatatg agagtataca catgtcatta gctatcaaaa ctatgatctg ttcagtaaac 2630

gttgtcca 2638

<210> 9

<211> 805

<212> PRT

<213> Mouse

<400> 9

Met	Ser	Ser	Ser	Ser	Trp	Leu	Leu	Leu	Ser	Leu	Val	Ala	Val	Thr	Thr
1				5					10					15	
Ala	Gln	Ser	Leu	Thr	Glu	Glu	Asn	Ala	Lys	Thr	Phe	Leu	Asn	Asn	Phe
			20					25					30		
Asn	Gln	Glu	Ala	Glu	Asp	Leu	Ser	Tyr	Gln	Ser	Ser	Leu	Ala	Ser	Trp
		35					40					45			
Asn	Tyr	Asn	Thr	Asn	Ile	Thr	Glu	Glu	Asn	Ala	Gln	Lys	Met	Ser	Glu
	50				55					60					
Ala	Ala	Ala	Lys	Trp	Ser	Ala	Phe	Tyr	Glu	Glu	Gln	Ser	Lys	Thr	Ala
65				70					75					80	
Gln	Ser	Phe	Ser	Leu	Gln	Glu	Ile	Gln	Thr	Pro	Ile	Ile	Lys	Arg	Gln
			85					90					95		
Leu	Gln	Ala	Leu	Gln	Gln	Ser	Gly	Ser	Ser	Ala	Leu	Ser	Ala	Asp	Lys
			100					105					110		
Asn	Lys	Gln	Leu	Asn	Thr	Ile	Leu	Asn	Thr	Met	Ser	Thr	Ile	Tyr	Ser
	115				120						125				
Thr	Gly	Lys	Val	Cys	Asn	Pro	Arg	Asn	Pro	Gln	Glu	Cys	Leu	Leu	Leu
	130				135					140					
Glu	Pro	Gly	Leu	Asp	Glu	Ile	Met	Ala	Thr	Ser	Thr	Asp	Tyr	Asn	Ser
145				150						155				160	
Arg	Leu	Trp	Ala	Trp	Glu	Gly	Trp	Arg	Ala	Glu	Val	Gly	Lys	Gln	Leu
			165					170					175		
Arg	Pro	Leu	Tyr	Glu	Glu	Tyr	Val	Val	Leu	Lys	Asn	Glu	Met	Ala	Arg
		180					185					190			
Ala	Asn	Asn	Tyr	Asn	Asp	Tyr	Gly	Asp	Tyr	Trp	Arg	Gly	Asp	Tyr	Glu
	195					200						205			
Ala	Glu	Gly	Ala	Asp	Gly	Tyr	Asn	Tyr	Asn	Arg	Asn	Gln	Leu	Ile	Glu
	210				215						220				

Asp	Val	Glu	Arg	Thr	Phe	Ala	Glu	Ile	Lys	Pro	Leu	Tyr	Glu	His	Leu
225					230					235				240	
His	Ala	Tyr	Val	Arg	Arg	Lys	Leu	Met	Asp	Thr	Tyr	Pro	Ser	Tyr	Ile

245 250 255  
 Ser Pro Thr Gly Cys Leu Pro Ala His Leu Leu Gly Asp Met Trp Gly  
 260 265 270  
 Arg Phe Trp Thr Asn Leu Tyr Pro Leu Thr Val Pro Phe Ala Gln Lys  
 275 280 285  
 Pro Asn Ile Asp Val Thr Asp Ala Met Met Asn Gln Gly Trp Asp Ala  
 290 295 300  
 Glu Arg Ile Phe Gln Glu Ala Glu Lys Phe Phe Val Ser Val Gly Leu  
 305 310 315 320  
 Pro His Met Thr Gln Gly Phe Trp Ala Asn Ser Met Leu Thr Glu Pro  
 325 330 335  
 Ala Asp Gly Arg Lys Val Val Cys His Pro Thr Ala Trp Asp Leu Gly  
 340 345 350  
 His Gly Asp Phe Arg Ile Lys Met Cys Thr Lys Val Thr Met Asp Asn  
 355 360 365  
 Phe Leu Thr Ala His His Glu Met Gly His Ile Gln Tyr Asp Met Ala  
 370 375 380  
 Tyr Ala Arg Gln Pro Phe Leu Leu Arg Asn Gly Ala Asn Glu Gly Phe  
 385 390 395 400  
 His Glu Ala Val Gly Glu Ile Met Ser Leu Ser Ala Ala Thr Pro Lys  
 405 410 415  
 His Leu Lys Ser Ile Gly Leu Leu Pro Ser Asp Phe Gln Glu Asp Ser  
 420 425 430  
 Glu Thr Glu Ile Asn Phe Leu Leu Lys Gln Ala Leu Thr Ile Val Gly  
 435 440 445  
 Thr Leu Pro Phe Thr Tyr Met Leu Glu Lys Trp Arg Trp Met Val Phe  
 450 455 460  
 Arg Gly Glu Ile Pro Lys Glu Gln Trp Met Lys Lys Trp Trp Glu Met  
 465 470 475 480  
 Lys Arg Glu Ile Val Gly Val Val Glu Pro Leu Pro Arg Asp Glu Thr  
 485 490 495  
 Tyr Cys Asp Pro Ala Ser Leu Phe His Val Ser Asn Asp Tyr Ser Phe  
 500 505 510  
 Ile Arg Tyr Tyr Thr Arg Thr Ile Tyr Gln Phe Gln Phe Gln Glu Ala  
 515 520 525  
 Leu Cys Gln Ala Ala Lys Tyr Asn Gly Ser Leu His Lys Cys Asp Ile  
 530 535 540  
 Ser Asn Ser Thr Glu Ala Gly Gln Lys Leu Leu Lys Met Leu Ser Leu  
 545 550 555 560  
 Gly Asn Ser Glu Pro Trp Thr Glu Ala Leu Glu Asn Val Val Gly Ala  
 565 570 575  
 Arg Asn Met Asp Val Lys Pro Leu Leu Asn Tyr Phe Gln Pro Leu Phe  
 580 585 590  
 Asp Trp Leu Lys Glu Gln Asn Arg Asn Ser Phe Val Gly Trp Asn Thr

26

595 600 605  
 Glu Trp Ser Pro Tyr Ala Asp Gln Ser Ile Lys Val Arg Ile Ser Leu  
 610 615 620  
 Lys Ser Ala Leu Gly Ala Asn Ala Tyr Glu Trp Thr Asn Asn Glu Met  
 625 630 635 640  
 Phe Leu Phe Arg Ser Ser Val Ala Tyr Ala Met Arg Lys Tyr Ser Ser  
 645 650 655  
 Ile Ile Lys Asn Gln Thr Val Pro Phe Leu Glu Glu Asp Val Arg Val  
 660 665 670  
 Ser Asp Leu Lys Pro Arg Val Ser Phe Tyr Phe Phe Val Thr Ser Pro  
 675 680 685  
 Gln Asn Val Ser Asp Val Ile Pro Arg Ser Glu Val Glu Asp Ala Ile  
 690 695 700  
 Arg Met Ser Arg Gly Arg Ile Asn Asp Val Phe Gly Leu Asn Asp Asn  
 705 710 715 720  
 Ser Leu Glu Phe Leu Gly Ile His Pro Thr Leu Glu Pro Pro Tyr Gln  
 725 730 735  
 Pro Pro Val Thr Ile Trp Leu Ile Ile Phe Gly Val Val Met Ala Leu  
 740 745 750  
 Val Val Val Gly Ile Ile Ile Leu Ile Val Thr Gly Ile Lys Gly Arg  
 755 760 765  
 Lys Lys Lys Asn Glu Thr Lys Arg Glu Glu Asn Pro Tyr Asp Ser Met  
 770 775 780  
 Asp Ile Gly Lys Gly Glu Ser Asn Ala Gly Phe Gln Asn Ser Asp Asp  
 785 790 795 800  
 Ala Gln Thr Ser Phe  
 805

<210> 10  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> PCR primer.

<400> 10  
 gactccgatc atcaagcgtc aacta

25

<210> 11  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>



<223> PCR primer.

<400> 11  
ggcagggagg catccagtgg

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 00/11932

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> IPC 7 C12N9/48 C12N15/57 C07K16/40 C12N15/10 C12Q1/68 G01N33/53 A61K38/48		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) IPC 7 C12N		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EPO-Internal, WPI Data, PAJ, BIOSIS		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL NUCLEOTIDE AND PROTEIN SEQUENCES, 29 June 1998 (1998-06-29), XP002147140 HINXTON, GB AC = AI036772. Sugano mouse kidney mklia Mus musculus cDNA clone IMAGE:1432919 5'clone similar to SW:ACE_DROME Q10714 ANGIOTENSIN-CONVERTING ENZYME PRECURSOR, SOMATIC ;mRNA sequence. EST abstract  <div style="text-align: center;">-/-</div>	10
<div style="display: flex; justify-content: space-between;"> <span><input checked="" type="checkbox"/> Further documents are listed in the continuation of box C.</span> <span><input checked="" type="checkbox"/> Patent family members are listed in annex.</span> </div>		
* Special categories of cited documents : <div style="display: flex; justify-content: space-between;"> <div style="width: 45%;">           "A" document defining the general state of the art which is not considered to be of particular relevance            "E" earlier document but published on or after the international filing date            "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)            "O" document referring to an oral disclosure, use, exhibition or other means            "P" document published prior to the international filing date but later than the priority date claimed         </div> <div style="width: 45%;">           "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention            "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone            "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.            "A" document member of the same patent family         </div> </div>		
Date of the actual completion of the international search  <div style="text-align: center; font-weight: bold;">13 September 2000</div>		Date of mailing of the international search report  <div style="text-align: center; font-weight: bold;">02/10/2000</div>
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016		Authorized officer  <div style="text-align: center; font-weight: bold;">Mateo Rosell, A.M.</div>

# INTERNATIONAL SEARCH REPORT

Int. Patent Application No  
PCT/US 00/11932

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	SOUBRIER F ET AL: "TWO PUTATIVE ACTIVE CENTERS IN HUMAN ANGIOTENSIN I-CONVERTING ENZYME REVEALED BY MOLECULAR CLONING" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA,US,NATIONAL ACADEMY OF SCIENCE. WASHINGTON, vol. 85, no. 24, 1 December 1988 (1988-12-01), pages 9386-9390, XP000113797 ISSN: 0027-8424 the whole document	1,2,5, 7-14,17
A	EHLERS M R W ET AL: "MOLECULAR CLONING OF HUMAN TESTICULAR ANGIOTENSIN-CONVERTING ENZYME: THE TESTIC ISOZYME IS IDENTICAL TO THE C-TERMINAL HALF OF ENDOTHELIAL ANGIOTENSIN-CONVERTING ENZYME" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA,US,NATIONAL ACADEMY OF SCIENCE. WASHINGTON, vol. 86, no. 20, 1 October 1989 (1989-10-01), pages 7741-7745, XP000113802 ISSN: 0027-8424 the whole document	1,2,5, 7-14,17
A	TAKEUCHI K ET AL: "PURIFICATION OF HUMAN LUNG ANGIOTENSIN-CONVERTING ENZYME BY HIGH-PERFORMANCE LIQUID CHROMATOGRAPHY: PROPERTIES AND N-TERMINAL AMINO ACID SEQUENCE" JOURNAL OF BIOCHEMISTRY,JP,JAPANESE BIOCHEMICAL SOCIETY, TOKYO, vol. 106, no. 3, 1 September 1989 (1989-09-01), pages 442-445, XP000051813 ISSN: 0021-924X the whole document	1,2,5, 7-10
A	WO 90 03435 A (INST NAT SANTE RECH MED) 5 April 1990 (1990-04-05)  fig. 3 the whole document	1,2,5, 7-15,17, 18
A	WO 91 00354 A (INST NAT SANTE RECH MED) 10 January 1991 (1991-01-10)  fig. 1 the whole document	1,2,5, 7-15,17, 18
	-/-	

# INTERNATIONAL SEARCH REPORT

Int. Patent Application No.  
PCT/US 00/11932

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 98 31835 A (UNIV MARYLAND ;HAGBERG JAMES M (US); UNIV PITTSBURGH (US); FERRELL) 23 July 1998 (1998-07-23) SEQ.ID.N.8 page 13, line 1 -page 20, line 25	1,2,5, 7-15,17, 18
A	WO 97 48801 A (GENETICS INST) 24 December 1997 (1997-12-24) SEQ.ID.N.14	10
A	TURNER A J ET AL: "MAMMALIAN MEMBRANE METALLOPEPTIDASES: NEP, ECE, KELL, AND PEX" FASEB JOURNAL,US,FED. OF AMERICAN SOC. FOR EXPERIMENTAL BIOLOGY, BETHESDA, MD, vol. 11, 1997, pages 355-364, XP000906774 ISSN: 0892-6638 the whole document	6
P,X	DATABASE EMBL NUCLEOTIDE AND PROTEIN SEQUENCES, 30 August 1999 (1999-08-30), XP002143332 HINXTON, GB AC = AL110224. Homo sapiens mRNA; cDNA DKFZp434A014 (from clone DKFZp434A014); partial cds abstract	10
P,X	DATABASE EMBL NUCLEOTIDE AND PROTEIN SEQUENCES, 26 December 1999 (1999-12-26), XP002147141 HINXTON, GB AC = AW258861. Sugano mouse kidney mKia Mus musculus cDNA clone IMAGE:2300959 5' similar to SW:ACE_RABIT P12822 ANGIOTENSIN-CONVERTING ENZYME PRECURSOR SOMATIC ; mRNA sequence. EST. abstract	10
P,X	WO 00 18899 A (MILLENNIUM PHARM INC) 6 April 2000 (2000-04-06) the whole document	1-21
P,X	EP 0 974 664 A (SMITHKLINE BEECHAM PLC) 26 January 2000 (2000-01-26) the whole document	1-21
E	WO 00 31278 A (ZYMOGENETICS INC) 2 June 2000 (2000-06-02) the whole document	1-21

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 00/11932

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9003435 A	05-04-1990	FR 2636968 A AT 121457 T DE 68922297 D DE 68922297 T EP 0388466 A JP 2854062 B JP 3502643 T US 5480793 A US 5359045 A	30-03-1990 15-05-1995 24-05-1995 14-09-1995 26-09-1990 03-02-1999 20-06-1991 02-01-1996 25-10-1994
WO 9100354 A	10-01-1991	FR 2649412 A CA 2035888 A EP 0432254 A JP 4500612 T US 5801040 A	11-01-1991 06-01-1991 19-06-1991 06-02-1992 01-09-1998
WO 9831835 A	23-07-1998	AU 6012798 A EP 0966543 A	07-08-1998 29-12-1999
WO 9748801 A	24-12-1997	US 5807703 A AU 3312197 A AU 3398697 A AU 3398797 A AU 3491497 A CA 2257827 A CA 2257834 A CA 2257842 A CA 2257847 A EP 0910637 A EP 0914429 A EP 0920499 A EP 0914430 A WO 9748798 A WO 9748799 A WO 9748800 A US 5807709 A US 5827688 A US 5837490 A	15-09-1998 07-01-1998 07-01-1998 07-01-1998 07-01-1998 24-12-1997 24-12-1997 24-12-1997 24-12-1997 28-04-1999 12-05-1999 09-06-1999 12-05-1999 24-12-1997 24-12-1997 24-12-1997 15-09-1998 27-10-1998 17-11-1998
WO 0018899 A	06-04-2000	AU 1311300 A	17-04-2000
EP 0974664 A	26-01-2000	JP 11318472 A	24-11-1999
WO 0031278 A	02-06-2000	AU 2150000 A	13-06-2000

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ BLACK BORDERS
- ☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☒ SKEWED/SLANTED IMAGES
- ☒ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☒ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**